DOCUMENT RESUME

ED 362 408 SE 053 699

AUTHOR	Pearsall, Marcia K., Ed.
TITLE	Relevant Research. Scope, Sequence, and Coordination of Secondary School Science. Vol. II.
INSTITUTION	National Science Teachers Association, Washington, D.C.
SPONS AGENCY	Department of Education, Washington, DC.; National Science Foundation, Washington, D.C.
REPORT NO	ISBN-0-873355-111-7
PUB DATE	92
CONTRACT	NSF-TPE-9050053
NOTE	269p.
AVAILABLE FROM	National Science Teachers Association, 1742 Connecticut Ave., NW, Washington, DC 20009 (Stock No. PB-97).
PUB TYPE	Guides - Classroom Use - Teaching Guides (For Teacher) (052) -- Books (010) -- Collected Works - General (020)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Ability Grouping; Academic Achievement; Classroom Research; Concept Formation; *Constructivism (Learning); *Curriculum Development; Curriculum Research; Educational Philosophy; Experiential Learning; Females; High Schools; High School Students; Inquiry; Minority Groups; *Misconceptions; *Science Curriculum; *Science Education; Scientific Concepts; Secondary School Science; Thinking Skills
IDENTIFIERS	Science Technology and Society Courses; Scope and Sequence; *Scope Sequence and Coordination

ABSTRACT

This book is a collection of research papers and respected philosophical statements on how secondary school students learn science best. Taken together, these papers serve as the basis for the Project on Scope, Sequence, and Coordination of Secondary School Science (SS&C). This book explains where the learning ideas of the SS&C Content Core came from and provides some evidence for how incorporating these ideas can improve students' knowledge, understanding, and appreciation of science. In the first section of this book the SS&C project is defined. The second section contains essays on the following SS&C Learning Statements: (1) Spacing Content Improves Learning--"The Spacing Effect" (Frank N. Dempster); (2) Learning Is Not Improved by Ability Grouping--"Achievement Effects of Ability Grouping in Secondary Schools" (Robert E. Slavin) and "The What, Why, and How of Cooperative Learning" (M. Lee Manning and Robert Lucking); (3) All Students Can and Should Learn Science--Encouraging Equitable Classrooms--"Multiple Inequalities" (Jeannie Oakes) and "Why Girls Don't Know" (Jane Butler Kahle); (4) Students Learn By Practice--Solving Problems, Designing and Carrying Out Experiments--"Education and Learning To Think" (Lauren B. Resnick) and "Science Education: It's Not Enough to 'Do' or 'Relate'" (Kathleen J. Roth); and (5) Learning Is Connecting New Information to Prior Knowledge--"Questions and Answers about Radical Constructivism" (Ernst von Glasersfeld), "Constructivism and the Empiricist Legacy" (Michael R. Matthews), "How Do Students' Views of Science Influence Knowledge Integration?" (Nancy Butler Songer and Marcia C. Linn), "Strategic Teaching in Science" (Charles W. Anderson), "Teaching Science for Understanding" (James A. Minstrell) and "Accommodation of a Scientific Conception" (George J. Posner, Kenneth A. Strike, Peter W. Hewson, and William A. Gertzog). (PR)

ED 362 408

SCOPE, SEQUENCE, AND COORDINATION OF SECONDARY SCHOOL SCIENCE

VOLUME II

RELEVANT RESEARCH

SE053 699

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Andrew Saindon

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

THE NATIONAL SCIENCE TEACHERS ASSOCIATION

BEST COPY AVAILABLE

VOLUME II
RELEVANT RESEARCH

SCOPE, SEQUENCE, AND
COORDINATION OF
SECONDARY SCHOOL SCIENCE

VOLUME II
RELEVANT RESEARCH

Edited by Marcia K. Pearsall

A PROJECT OF
THE NATIONAL SCIENCE TEACHERS ASSOCIATION
1992

Copyright © 1992 by The National Science Teachers Association

The National Science Teachers Association
Scope, Sequence and Coordination of Secondary School Science
1742 Connecticut Ave., NW
Washington, DC 20009

This material is based upon work supported by the National Science Foundation under Grant No. TPE-9050053.

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Library of Congress Card Catalog Number: 92-60225
ISBN: 0-873355-111-7
NSTA Stock Number: PB-97

Cover and book design by Marcia K. Pearsall
Cover photo © The Image Bank, Ilene Astrahan

Printed in the United States of America
10 9 8 7 6 5 4 3 2

CONTENTS

Acknowledgments iii
Introduction v

DEFINING THE PROJECT

Essential Changes in Secondary Science
Bill G. Aldridge 1

Project on Scope, Sequence and Coordination
Bill G. Aldridge 9

SS&C LEARNING STATEMENTS

Spacing Content Improves Learning

The Spacing Effect
Frank N. Dempster 25

Learning is Not Improved by Ability Grouping

Achievement Effects of Ability Grouping in Secondary Schools
Robert E. Slavin 43

The What, Why, and How of Cooperative Learning
M. Lee Manning and Robert Lucking 69

All Students Can and Should Learn Science---Encouraging Equitable Classrooms

Multiplying Inequalities
Jeannie Oakes 79

Why Girls Don't Know
Jane Butler Kahle 111

Students Learn by Practice---Solving Problems, Designing and Carrying Out Experiments

Education and Learning to Think
Lauren B. Resnick 129

Science Education: It's Not Enough to 'Do' or 'Relate'
Kathleen J. Roth 151

Learning is Connecting New Information to Prior Knowledge

Questions and Answers about Radical Constructivism
Ernst von Glasersfeld 169

Constructivism and the Empiricist Legacy
Michael R. Matthews 183

How Do Students' Views of Science Influence Knowledge Integration?
Nancy Butler Songer and Marcia C. Linn 197

Strategic Teaching in Science
Charles W. Anderson 221

Teaching Science for Understanding
James A. Minstrell 237

Accommodation of a Scientific Conception
*George J. Posner, Kenneth A. Strike, Peter W. Hewson,
and William A. Gertzog* 253

Index

ACKNOWLEDGMENTS

The project would like to express sincere thanks to the following individuals who offered constructive criticism on the draft of this document: Richard A. Duschl, University of Pittsburgh; William C. Kyle, Jr., Purdue University; Robert Slavin, Center for Research on Effective Schooling for Disadvantaged Students; and Emmett L. Wright, Kansas State University.

We also appreciate the work of Ronald J. Bonnstetter, Director of Secondary Science Education at the University of Nebraska, who coordinated the review effort.

We acknowledge the Department of Education and the National Science Foundation for providing the generous financial support for the Project on Scope, Sequence and Coordination of Secondary School Science.

Finally, the dedicated work of our pilot centers merits special recognition as they serve as the leaders in the science education reform effort across the country.

California Department of Education
Sacramento, CA
Thomas Sachse
916/657-2914

Anchorage School District
Anchorage, AK
Emma Walton
907/269-2274

University of Puerto Rico
San Juan, Puerto Rico
Manuel Gomez
809/764-8369
Anacet Bernacet
809/765-5170

University of Iowa
Iowa City, IA
Robert Yager
319/335-1189

University of North Carolina at Wilmington
Wilmington, NC
David Andrews
919/395-3887
East Carolina University
Greenville, NC
Charles Coble
919/757-6172

Baylor College of Medicine
Houston, TX
Linda Crow
713/798-4613

INTRODUCTION

What is this book about?

This book is a collection of research papers and respected philosophical statements on how secondary school students learn science best. Taken together, these papers serve as the basis for the Project on Scope, Sequence, and Coordination of Secondary School Science (SS&C).

This book, ***Volume II: Relevant Research****, is a companion volume to the curriculum guide produced by the SS&C project in March 1992:* ***The Content Core****.* ***The Content Core*** *described the tenets of scope, sequence, and coordination and then presented the subject matter of science organized according to those tenets. The earlier volume asserted the desirability of science curricula that incorporated certain learning ideas, such as "spacing content improves learning." This second volume spells out where the learning ideas came from and provides some evidence for how incorporating these ideas can improve students' knowledge, understanding, and appreciation of science. Thus the primary aim of* ***Volume II*** *is to present respected expressions or evidence of the learning ideas that comprise SS&C.*

It is our hope that this collection of research papers will achieve multiple ends. First, it will illuminate the basis of the project—make evident the different educational, cognitive, and philosophical strands concerning science education that were/are woven together into the "fabric" of SS&C. We indicate in most chapter introductions how each selection relates to SS&C.

Second, the collection will serve as a science education resource. Individual articles define or elaborate key terms in science education such as "constructivism" or "higher order skills." Almost every selection has its own bibliography that directs interested persons to additional readings. We encourage educators to use this volume as a starting point for their efforts to improve science education in the secondary schools. We believe that when teachers have direct access to science education research, they can make informed judgements ***for themselves*** *as to whether the results are relevant to their own educational circumstances. They can determine whether proposed modifications could positively impact their students' learning/understanding of science.*

How were the articles selected?

The bibliography of ***The Content Core*** *was the starting point for this collection. Bill Aldridge, the founder of the SS&C project, frequently cited these works as the basis or support for SS&C project tenets. Project staff used additional works when implementing SS&C in the pilot centers and advocating its use in secondary schools. Finally, researchers in science teaching were consulted to review the references and recommend additional, current research. It should be noted that this volume is not a defini-*

tive collection of all current research and trends in science teaching, but rather it is a gathering together of a few, select papers to provide an interesting, accessible introduction to the learning ideas that comprise SS&C.

Why should practicing science teachers read or keep up with this research?

In the first volume of NSTA's "What Research Says to the Science Teacher," Patricia Blosser, then a faculty research associate in science education at Ohio State University, addressed this issue. Blosser simply stated, "Awareness must precede action." In other words, if teachers know how students learn science best or the teaching strategies that have been found most successful in today's science classrooms, they can then adapt or modify their classroom practices accordingly. If the teachers have the information in their hands, they can shape the design of curricula or courses because they possess the necessary knowledge base. We believe that if restructuring efforts are to succeed, the teachers must be active participants in the planning and implementation process. We hope that this volume will inspire teachers to reflect on how science teaching and learning occur in their classrooms, to explore alternatives or modifications, and to request, support, and participate in restructuring efforts.

What are the uses of this book?

This question is answered in part by the previous questions. We believe that this book can help teachers who are implementing (or just interested in) SS&C to become acquainted with the learning ideas on which it is based. But we also believe that this book can be used to "speak to" educational policy makers. With these research papers in hand, science teachers and supervisors can make convincing and persuasive arguments about the need for restructuring science curricula. Additionally, this ***Volume II*** *can be a teaching tool—either for teacher inservice or preservice programs. This book easily translates to a syllabus. Finally, we hope that this book inspires further research. Many authors included note where there is simply a lack of information or evidence about how students learn science best. Both the classroom teacher and the educational researcher could pursue these areas open for investigation.*

How is this book organized?

The book is organized into two sections. The first section defines the SS&C project using the two key papers written by project founder and NSTA Executive Director Bill Aldridge. The second section proposes five learning "statements" that, taken together, undergird the project Aldridge proposed. Following a brief introduction to the learning idea, several excerpts or papers are presented. These selections range in style and format from a respected philosopher's comments at a conference to a technical report of a psychology experiment. We feel that the diversity of the documents represents the varied perspectives concerning the learning and teaching of science. Each article is reprinted here with full documentation and with special attention to the integrity of the original. The only exceptions are the few selections that were edited due to their length, and they are so identified.

What happens next with the SS&C project?

Along with ***The Content Core****, this volume serves as the second effort to flesh out the ideas of SS&C. A third component, and by far the most important, is the actual experiences of our pilot centers as they implement and evaluate SS&C. Pilot centers* *are now established in California, Texas, Iowa, North Carolina, Alaska, and Puerto Rico. Each center strives to provide science instruction in four subjects every year---physics, chemistry, biology, and earth/space science---and to every student in grades 7-12. In the middle level grades, some programs use integrated curricula; some use STS curricula; and still others retain the separate subjects. At the high school level, students encounter the subjects as separate disciplines but with strong coordination.*

From time to time, newsletters from the pilot centers and the NSTA Coordinating Center report on program and classroom experiences. Presentations, progress reports, and panels are given at regional and national meetings. We intend to report on SS&C in the field in a more comprehensive way by producing a ***Volume III: Implementing SS&C****, probably in Fall 1993. This third volume not only will report on the utility of the first two volumes, but provide specific accounts of how the pilot centers have dealt with such issues as instructional materials, assessment, teacher training, and institutional support. Any educational reform, theory and research nonwithstanding, will succeed or fail only by the test of how well it works with students.*

DEFINING THE PROJECT

BEST COPY AVAILABLE

DEFINING THE PROJECT

The two articles in this section, both written by Bill G. Aldridge, present the initial tenets and evolving objectives of the Project on Scope, Sequence and Coordination of Secondary School Science. Aldridge, Executive Director of the National Science Teachers Association, calls for the reform of science education in the United States. His two articles provide the critical framework under which the subsequent articles in this volume are collected. Articles were selected according to their ability to inform or engage specific principles undergirding the SS&C project.

*The first Aldridge article is the founding document of the SS&C project. Published in the January/February 1989 issue of **NSTA Reports!**, the article invited science teachers to participate in an educational reform project and generated intense interest from educators across the United States. In the article, Aldridge identifies three key areas for science education reform which he names "scope," "sequence," and "coordination." Primarily by making comparisons to the science curricula of the People's Republic of China and the then-Soviet Union, Aldridge proposes modifications to American science programs. These modifications include the study of science every year for the six secondary years, the coordination of the four traditional disciplines, and the revisiting of basic concepts on a periodic basis (spaced learning). Drawing on the work of researchers in learning, Aldridge also recommends that students first encounter science content on a descriptive and phenomenological level and then progress through the empirical and semi-quantitative to the theoretical and abstract levels.*

The second Aldridge article elaborates the earlier reform proposal and documents the structuring and funding of the actual SS&C project. Aldridge indicates that the project will synthesize the best of research and development efforts on learning and, in some detail, he argues that a focus on the questions, "What do we mean?" "How do we know?" and "Why do we believe?" is a necessary component of the reform. He briefly discusses current research in preconceptions (or misconceptions) and student motivation. Finally, he describes a new technology to be tested in the project—the Compact Disc Interactive—and its potential as a tool of performance-based assessment.

ESSENTIAL CHANGES IN SECONDARY SCIENCE: SCOPE, SEQUENCE, AND COORDINATION

Bill G. Aldridge

Secondary school science courses in the United States are major filters to careers in science or science-related fields. With increases in science requirements for graduation from high school in almost all states, secondary school science courses have gained the additional reputation of contributing to the already high dropout rates between grades 9 and 12.

A careful analysis of the existing scope, sequence, and coordination of science subject matter in U.S. secondary schools reveals very serious deficiencies. Needed changes in secondary school science would greatly increase the numbers of children who study science and would produce greater numbers of persons entering scientific and engineering careers. (At present, these represent about five percent of all students.) The proportions of underrepresented groups in these careers would also increase. At the same time, the remaining 95 percent of the students would become sufficiently literate to participate effectively in our scientific and technological world.

The changes needed are, in part, suggested by analyses of science programs in other nations, particularly in the USSR and the People's Republic of China. What are these changes, and how could they be implemented in the United States?

A recent SRI International (formerly Stanford Research Institute) report to the National Science Foundation suggests that the NSF support an educational model that will bring a larger population into science careers, thus "broadening the pool" rather than "skimming the cream."

The present sequencing of courses and tracking of students reinforce the accumulation of advantage for some students and effectively preclude others from later selecting science as a career. A child who has not enrolled in algebra in the eighth or ninth grade and then not taken the other more advanced math courses in sequence is not ready for the series of advanced science courses required in high school. As a result, a student is ruled out from majoring in one of the physical sciences or engineering fields in college.

---

*This article first appeared in **NSTA Reports!** (January 1989), pp. 1, 4-5, Washington, DC: National Science Teachers Association.*

High school physics, chemistry, and biology courses duplicate substantial portions of introductory college courses. Young people who take these high school courses enter college with prior knowledge of the subjects, and they also score higher on entrance exams. They therefore are perceived as having higher ability. They are given scholarships and opportunities to study under the best faculty, so that their prior advantages are rewarded, leading them to continue in these science fields.

Young people who did not take these courses in high school, but later show interest, are at such a disadvantage that without special assistance or extraordinary motivation and persistence, they will never be able to compete successfully with their more advantaged peers. Since such disadvantage in the physical sciences is often associated with women and minorities, it is not surprising that we find their numbers disproportionately low among engineers and physical scientists.

How then can we encourage the child who has not yet achieved significantly in science or math? We know that motivation is well established as a major variable in learning. Prior success strongly affects students' self-esteem and self-image. Many children are led incorrectly to believe that they lack the ability to learn science or math when they compare their achievements with those of their more experienced and advantaged peers. They avoid studying science and math because they believe they lack some special ability to succeed in these subjects.

Yet most of the achievement of students who do complete the secondary school science courses is measured by how well they acquire information and facts. Few secondary school students understand science in terms of "What is the evidence?" "How do we know?" "Why do we believe?" and "How do we find out?" If we focused secondary school science more on these skills than on the acquisition of information, all students would achieve better. We would also see a different distribution of children among the high achievers. Learning what, how, and why is pleasurable and highly motivating, and relies far less on prior knowledge or advantage. These are the elements of scientific literacy.

Recent NSF studies, as well as most other major reports, urge us to focus on science education for the general public. We should concentrate our resources and attention on the majority of students rather than upon the few advantaged children. Motivating the so-called "average" student to take more science courses will ultimately lead to a totally different mix among outstanding scientists than we have at present.

The following analogy based on molecular motion is appropriate to our problems in science teaching. Imagine a container of gas molecules of several kinds moving in all directions at various speeds. Some that move slowly now will, when struck, move much more rapidly. Suppose you want to select the fastest molecules moving in a particular direction. You could place a small opening in the container, and when a fast molecule approaches the opening, push the molecule in the direction you want it to go. In this way you can select the fastest molecules. You would, of course, be ignoring even faster molecules moving in other directions. You would also be ignoring the fact that other molecules will change speed at other times and may be moving even more rapidly than the ones you selected.

The present state of secondary school science is like the first state of the gas. We select the young people who happen at a certain time to be moving fastest in the "right" direction, and we ignore even faster (more able) children who do not happen at *that* *moment* to be moving in the "right" direction. We give those selected students the added push (advantages in the form of resources, teacher attention, and recognition), and we ignore the rest of the students.

There is another way of making fast molecules move toward the opening. If you heat the entire sample of molecules, all will move faster. The fastest will be even faster than before, and more molecules will pass through the opening. And you've accomplished this without selecting certain molecules or giving only them special assistance.

Heating the gas is analogous to providing more and better science education for *all* children so they can *all* move faster. Many of them will move in the "right" direction, but not necessarily at the time we want them to do so. If we allow for this delay, we will reap a rich selection of students in science, just as we got a greater variety of molecules in the heated gas.

How then do we make secondary school science accessible to all students? We must attack the major cause of disinterest and failure: the way science courses are structured, sequenced, and taught and the complete lack of coordination among them.

Consider the typical situation in the United States. Children study science in elementary school mainly by reading about it. In junior high, they study earth and space science, physical science, and biology—but not in an appropriate sequence and without coordination among the subjects. Great variations also exist from one school and school district to another, with no accommodation made for the high mobility rate of U.S. families.

At the senior high school level, the three major science subjects are biology, chemistry, and physics. Geology, earth science, astronomy, other life science courses, and myriad related subjects are offered as well. But NSTA surveys show that the vast majority of high school students who enroll in science take biology, chemistry, or physics.

High school courses in biology, chemistry, and physics are very similar in scope and sequence to their college-level counterparts. Their approach is somewhat lower in quantitative and theoretical terms, but much overlap exists. Both students and teachers see high school science courses as preparation for college science courses. Few students take all three subjects unless they are required to do so or they intend to major in science or engineering in college.

Almost no general students take all three high school science courses because they perceive the subjects as tough and rigorous and believe that only "able" or "college-bound" students are likely to succeed in them. Most high school students simply do not believe they have the ability to learn this material. Most teachers and counselors agree and directly or indirectly convey this opinion to their students.

The recent trend toward raising science and math requirements for graduation from high school has created a severe problem for both teachers and students. To graduate, students must now take at least two science courses. In some states they must take three. What has happened?

In some cases the two- or three-course science requirement is sparking creative responses as schools add new courses and use old courses in other areas as substitutes. The latter include vocational-technical courses and health or nutrition courses. In other instances, when faced with taking biology, chemistry, and physics, many students drop out, further increasing our tragically high national dropout rate of more than 25 percent.

BILL G. ALDRIDGE

The fundamental problem with high school biology, chemistry, and physics courses is that they are not coordinated, are highly abstract and theoretical, do not spend enough time on each subject, and do not use correct pedagogy. In short, we never give students the chance to *understand* science.

For example, high school physics students are, from the first day, exposed to the highest levels of abstract thinking, and this continues unrelentingly for the next 179 days of the school year. All of introductory physics is concentrated into that one year, and nothing in the students' experience has prepared them for this pace and level. Students do not have enough time to begin to assimilate even the basic concepts, theories, and applications. As a result, very few young people who take high school physics retain anything they learn even a short time after the course is over. Neither the student nor the teacher is at fault. The curriculum itself is the problem. We should not be surprised, therefore, that so few students take any more high school science than they are forced to, and most dislike the courses they do take.

By contrast, consider the secondary school-level science programs offered in the USSR and the People's Republic of China. In both countries, all—yes, *all*—students take several years of biology, chemistry, and physics. And essentially all children learn these subjects successfully. The dropout rates in these countries are only about 2 percent compared with our 25 percent.

Therefore, not only do these countries teach *more* science to *more* students, but they do so to more of those students we call "low ability" or "non-college-bound." How do they do it? They carefully design the scope and sequence of their science programs and make sure the subjects are well coordinated with one another. Each of the subjects is taught over several years.

As you can see in Table 1, the average student in the USSR and China spends almost twice as much time studying biology and chemistry and more than three times as much time studying physics than his or her American counterpart. Notice that the ratios are even higher when you compare the years spent on each subject.

The secret of both countries' success in teaching science is multifaceted. First, they offer the subjects over four or five years instead of just one year, sometimes devoting only one or two hours a week on a subject for one or more of those years. Second, they sequence the content and approach, starting with the concrete, phenomenological, and descriptive in the early years to the semi-quantitative and empirical in the middle years and then the theoretical and abstract in the later years.

When you analyze the content and level of the last year for each of these subjects, you find that they are almost identical with what we offer our students. But our students don't have the four or five years of prior grounding in experience and applications. No wonder so few of our students learn the subjects and why so many see our courses as hopelessly difficult.

In the USSR and the People's Republic of China, each high school science course includes many practical applications, so that young people understand why they need to learn science. Finally, these countries closely coordinate the content from one subject to another. All these reasons for their success in teaching science are based on what we know or should know from psychologists about how people learn.

Table 1

Time Spent on Biology, Chemistry, and Physics

	United States	**U.S.S.R.**	**P.R. of China**
Biology	180 hrs.	321 hrs.	256 hrs.
	1 year	6 years	4 yrs.
Chemistry	180 hrs.	323 hrs.	372 hrs.
	1 year	4 years	4 years
Physics	180 hrs.	492 hrs.	500 hrs.
	1 year (21%)	5 years	5 years

Here is an example of how the USSR coordinates its science. During a certain number of class hours in one year, students in biology class explore the human heart and the circulatory system. In chemistry class, the same students investigate chemical reactions that involve oxygen and, in particular, such metabolic reactions as oxygen transport by hemoglobin. At the same time in physics class they study the kinematics and dynamics of fluid flow. Now *that* is coordination!

For those who say that these countries lack breadth in their overall education, look again. Every child is also required to study government, geography, history, music and art, languages, and ship skills. The latter, which are required every year, are not like our dead-end vocational programs. Shop skills offer students essential experience working with their hands to make, shape, bond, and cut things—skills that give them the option to become an outstanding experimental scientist, a laboratory technician, or a handyperson in their own lives. This broad curriculum is possible simply by extending the school year and the school day.

Some may protest that such a system could never work here because we are a nation of 16,000 school districts. But who controls the curriculum now? Not the 16,000 school districts, but the textbook publishers who simply react, as they must, to the market. So let's change the market. The publishers will happily respond.

Let me offer two possible scenarios, illustrated in Tables 2 and 3, for redesigning our high school science curriculum, or more ambitiously, our entire secondary school science curriculum.

Notice in both Tables 2 and 3 that at first we would expose students most intensively to the most descriptive and phenomenological parts of the three sciences. The emphasis on the most abstract and theoretical would be saved for the later years. Empirical and semi-quantitative treatments would be emphasized in the middle years. Most important, students would be taught science in a way they would be able to understand and apply it, whether as scientists or citizens.

Tables 2 and 3 demonstrate only two *possible* scenarios for changing secondary school science. Coordination is essential if these various components are to be learned

Table 2

Proposed Example of a Revised Science Curriculum For Grades 10 Through 12

	Grades			Total Time Spent
	10	11	12	
Hours Per Week By Subject				
Biology	3	2	1	216
Chemistry	2	2	2	216
Physics	1	2	3	216
Total Hours Per Week	6	6	6	
Emphasis	descriptive; phenomenological	empirical; semi-abstract	theoretical; abstract	

Table 3

Proposed Example of a Revised Science Curriculum For Grades 10-12 in the United States

	Grade Level						Total Time Spent
	7	8	9	10	11	12	
Hours Per Week By Subject							
Biology	2	2	2	2	1	1	360
Chemistry	1	1	2	2	2	2	360
Physics	1	1	1	2	2	3	360
Earth Science/Geology	3	2	2	1	1	1	360
Space Science/Astronomy	1	2	2	2	2	1	360
Computer Science	1	1	2	2	2	2	360
Science/Technology/Society	2	2	1	1	2	2	360
Total Hours Per Week	11	11	12	12	12	12	
Emphasis	descriptive; phenomenological		empirical; semi-quantitative		theoretical; abstract		

properly. And, of course, any such broad changes in curricula would pose significant staffing, scheduling, and other problems. But change can only come from exploring what is ideal and possible and then working from there to what is practical.

When are we going to change secondary school science in this country? Science need not be the terrible filter that it now is. It need not be an unpleasant experience for so many. Why not start a movement now to create a sensible scope, sequence, and coordination of secondary school science? Let's try it in a school or district and see if we get the dramatic improvements in student attitude and achievement that I predict would occur.

Most people are able to learn physics, chemistry, biology, geology, and other related subjects. People can learn how these subjects apply to their lives and to society's problems. If we teach science to all students in a coordinated way from the concrete to the abstract, with practical applications, our future citizens will be more evidence-oriented. They will know how and when to ask questions, how to think critically, and then how to be able to make important decisions based on reason rather than on emotion or superstition.

Science made understandable and accessible to all will also mean more and better scientists and engineers and a different mix among them, including larger numbers of such now underrepresented groups as women and minorities.

PROJECT ON SCOPE, SEQUENCE, AND COORDINATION: A NEW SYNTHESIS FOR IMPROVING SCIENCE EDUCATION

Bill G. Aldridge

Background and Relevant Research

As numerous reports and studies indicate, schools in the United States are failing to educate students for a world that depends more and more on sophisticated and rapidly changing science and technology. Surveys show that the majority of students leave our schools without a basic understanding of science, mathematics, or technology. Neither the demand for scientists and engineers is being met, nor are schools preparing future citizens with an adequate background of knowledge necessary to make decisions affecting their lives.

Students perceive currently structured, textbook-driven science subjects as difficult, boring, and not relevant to their lives. Many opt out as soon as the system allows—over half our students don't take another science course after 10th grade. Only 19% of high school students take a course in physics and only about 40% take a course in chemistry.

In this setting, as it existed in the Fall of 1988, the Project on Scope, Sequence and Coordination of Secondary School Science (SS&C) germinated. The author had just concluded an analysis of science education in several countries and was struck by the obvious differences between science education in the U.S. compared with other countries (Wirszup, 1986; Szekely, 1987, Janke, 1983, 1987, 1988). For example, in essentially all other industrialized nations of the world, students study science in several subjects over several years. Only in the United States do we use the *layer cake* of biology, chemistry, and physics at the high school level. Most apparent was the fact that, while we tend to think other nations educate only the elite while we teach all students, it is really just the opposite. In this country, we teach science only to the elite, and we somehow assume that others cannot learn science.

---

*This article first appeared in **Journal of Science Education and Technology** (March 1992), Volume 1, Number 1, pp. 13-21. New York: Plenum Publishing Corporation. Reprinted by permission.*

There is an incorrect assumption widely held that only certain children are capable of learning science and mathematics. Students, identified at an early age as "most able" or "most intelligent," are tracked into future course work in science and math. These students are often the most advantaged. The remainder are systematically filtered out of science and math. But science is needed by everyone and everyone is capable of learning and enjoying science. Students should stay in science until they have had several years of a good learning experience in each of well-coordinated science subjects.

In these reports and other research efforts there was also clear progress being made in research on learning, particularly as it applied to the physical sciences. The work of Resnick (1987) and Driver (1984-89) are compelling. The complexity and difficulty of learning what scientists and science teachers generally regard as elementary science concepts are found to be much more difficult than is apparent on the surface. Preconceived metaphors and explanations for natural phenomena are found to persist even when the subject has been taught, and even when students are successful in examinations. Some of the research which had been done by Karplus (1977), Arons (1983, 1990), and McDermott (1984) was also of great interest. These results, too, show how little depth of student understanding is found when probing questions are asked about the science involved. Also, there emerges a general pattern of experience and sequence through which everyone must pass before abstractions can be understood.

The author, having lived through the curriculum projects of the 1960s and having worked at the National Science Foundation (NSF) in science education, was familiar with a range of materials development projects, most of which, unfortunately, seemed far removed from results of these and other promising research.

The SS&C project idea was to form a synthesis from the best research and development efforts from these and earlier research by Piaget (1973). The various research efforts, and their successes, had clear implications for reform. Research on the spacing effect (Dempster, 1988) implied that science should be taught a few hours per week over several years instead of concentrated into one year. The work of Bruner (1960) and others (Arons and Karplus, 1976) has shown the importance of sequencing. Piaget's work suggests that concrete experiences with science phenomena should precede terminology. Similarly, concepts should be derived from experience, with students acquiring a concept from experience with a phenomenon in different contexts. Once concepts are established, they should be symbolized and those symbols related to each other. These more complex relationships would be constructed over time.

A synthesis of this various research indicates that a framework should be built around three fundamental questions: What do we mean? How do we know? and Why do we believe?

What do we mean?

"The structure of the natural world requires language to explain it. But there are many ways to describe natural phenomena. Customarily, we recognize a phenomenon and give it a name...Sometimes in science curricula, remembering the names and their definitions seems to become an end in itself. *But a name should not become more important than the phenomenon being described, or than its empirical or logical relationships with other phenomena*" (from the California Framework).

Science uses many terms. These terms may be names for things, or they may be names for concepts, principles, laws, or theories. In writing science materials, it is very important to recognize that a term is a name for something. Giving the name does not explain it. For example, saying that something falls because of gravity tells us nothing. In this context, gravity is just a name.

When a name is for a specific thing, such as a person's name or the name of a cat or a dog, the name tells you nothing about the person, cat, or dog.

When one uses a scientific name, such as that for a plant or an element, the name has more generality. It means that every plant or element with these names is identical in some way. The name is agreed upon so that we can communicate about the plant or the element. But the name is still arbitrary, and by itself has no meaning. It has meaning only to those who understand which characteristics of the plant or element are inherent to the names we have agreed upon. And understanding those characteristics requires **experience** with the plant and element and with the characteristics used to categorize the plant or element.

A term is often defined so as to represent a concept in science. Examples include mass, speed, velocity, acceleration, cell, tissue, force, pressure, energy, gravity, and so on. Unless the person learning a term first has had **experience** with the phenomenon the term represents, **in its various contexts**, a word for the term is meaningless. Sometimes meaning can be elicited from past experience, or it can be created from vicarious experience. As experience occurs with the phenomenon in various contexts, the concept forms in the mind of the person having the experience. The definition of the term in words or symbols is quite meaningless without this prior experience.

Some concepts, such as force, pressure, energy, or acceleration, are derived from other concepts. For example, the concept of pressure requires prior acquisition of the concepts of force and surface area. If the latter two concepts are not part of the person's cognitive repertoire, then the concept of pressure cannot be developed, and the term has no meaning for that person.

There are other terms that are names for phenomena that occur in science. These terms should not be used until a person has experience with the phenomena. For example, there is a phenomena called refraction. One can observe situations in several contexts where light bends at the boundary between two different transparent substances. The name of this bending is called refraction. There are many terms that stand for phenomena: refraction, dispersion, reflection, freezing, boiling, cell division, etc. When phenomena occur and students are, for the first time, trying to connect a name with a phenomenon as a first step to understanding it, we must not immediately begin some sort of explanation of it. To do so would confuse a theory that accounts for the phenomenon with the existence and name for the phenomenon. Brownian motion is a phenomenon. It can be observed. But you do not immediately "explain it." Sound and light both exhibit various phenomena. But textbooks often impose a theory onto the phenomena by naming the observations using terms from a theory. They call sound "sound waves" or light "light waves" or "photons." Sound is sound and light is light. One needs to get at the phenomenon first; name it after experience with it; and only then try to examine the various models or theories [explanations] which can best account for that phenomenon.

How do we know?

This question concerns what we call science facts. There are many facts taught in science: Light travels at 3 X 10^8 m/s; The volume of a gas is proportional its absolute temperature when the pressure is held constant; Bacteria do not have nuclei; Red blood cells contain hemoglobin, which carries oxygen to tissues; Most atoms are made up of electrons, protons, and neutrons; The percent of oxygen by volume in air is 20.946%; The average distance from the Earth to the sun is 92,900,000 miles; The angle of reflection of a light ray equals its angle of incidence, and the incident and reflected rays are in the same plane; The Earth rotates on its axis and revolves around our sun; etc.

Facts are often measurements. Some measurements are complicated and difficult to make. Many science measurements are indirect, and require understanding of concepts, principles, and laws, just to carry out the measurements. Sometimes, special instruments are needed. But knowing a measurement fact is, itself, quite worthless; what is important is how we know. How was the measurement made? When we understand **how** we know, we have more confidence that the fact being stated is correct and we better understand its limitations.

Some facts are empirical laws. They are summaries of observations, or they are observed relationships between science concepts. Charles' Law is an empirical law. So is the law of reflection or the law of refraction (Snell's law). It is very important that students understand **how** we know an empirical law is valid, and, especially, under what conditions. Unsupported assertions of such laws is the worst kind of science teaching! The students should first have developed, through experience, the concepts that are related by the law, such as the concepts of volume and temperature (and of pressure, which is held constant) in Charles' Law. Then they should either conduct the investigation themselves, or read a careful description of the investigation (How we know). They then should examine the results and arrive at the relationship through their own analyses. When such empirical laws are encountered, they should not be immediately explained. This process imposes a model or theory onto the law. Words such as "this is because" are used, without the person using them even realizing that he or she is imposing some theory. For example, the empirical law called interference is often stated as an unsupported assertion, followed by the "explanation," "this is because light is a wave."

Why do we believe?

What do we mean? How do we know? These are questions involving meaning and facts. But science has more to it than objective fact. It also involves belief. Such beliefs occur for reasons, however. In some cases, we use different facts along with a process of logical deduction to arrive at a new fact. In doing so, we are expressing a belief in the process called logic, or logical deduction. We also create models or theories to account for facts, observations, and empirical laws. These can be visual models, or they can be abstract theories in the form of mathematical equations. Scientists believe in such theories only so long as they are successful in accounting for all aspects of the phenomenon. A theory or model is not only helpful in visualizing or expressing relations among various laws, but also it can often make new predictions of facts, phenomena, or observations. In fact, to change a theory, new predictions are required, along with confirmation by independent scientists.

In textbooks and other instructional materials, it is very important to have students understand **why** we believe in certain theories. Why do we believe that the **theory** of natural selection accounts for the **facts** of evolution? Why do we believe that there are atoms, molecules, electrons, protons, and neutrons? Why do we believe that they have the size and structure which they have? Why do we believe that atoms and molecules in matter are moving in random motion and that this motion reflects something else we call temperature? Why do we believe that light is electromagnetic and behaves like a wave? And in this connection, there must be much more than just some definition of these terms. Why do we believe that light and sound both behave like a wave? Why do we believe that light also behaves like a particle (photon)? And why is it that belief in both is in no way paradoxical, but instead reflects the reality that light is neither a wave nor a particle. Light is light, and there are theories we create to account for what light does in different circumstances. Why do we believe that the facts of continent movement on Earth, and of its ocean features and other physical data, are best accounted for by a theory called plate tectonics?

It is extremely important **not** to study these theories at all until students understand the observations, phenomena, concepts, facts, and empirical laws that the theories attempt to encompass. Models or theories should never be imposed at the point when facts, phenomena, or empirical laws are first introduced. Nor should they ever be used to introduce such facts, phenomena, or laws. As an example, temperature is often defined in textbooks as average kinetic energy of molecules. Not only is the definition wrong, it uses a theory, kinetic theory, to explain a basic concept, temperature. The temperature concept needs first to be acquired through experience, leading to understanding of its proper definition. The concept of a molecule needs to be developed; then, when the student understands empirical laws associated with solids, liquids, and gases, he or she can develop models leading to kinetic theory.

Qualitative relationships would be first; then measurements would be made and empirical relationships found; finally theories would be constructed from which new predictions could be made. Following this progression, science teachers would go from direct experience with phenomena to terms and concepts and then to laws, principles, and theories.

Spacing and sequencing facilitate optimal student learning, but they are not sufficient to achieve such learning. There is a prevalent belief that creating a well-designed **instructional** package leads to learning. Some believe that a "perfect" lecture, clear, organized, and eloquent, should teach students. Others believe that a well-illustrated exposition with clever or interesting presentations will do so. Many producers of educational technology believe that fancy simulations, beautiful graphics, full-motion video, or expensive films are what science education needs. However, evidence on learning shows that people **construct** their own knowledge (Driver, 1984-1989, Arons, 1976; Linn, 1986). It cannot be transmitted to them. And regardless of the quality of the presentation, this transmission mode will not produce the necessary kinds of learning.

It is important that we not merely offer clever or well-designed expositions of theories or models. No matter how clearly expressed, graphically or visually superb, or ingenious the explanation, it is still just an explanation! We know that students learn by creating their own models and theories, not by reading or studying those developed by scientists and presented in your text. Before reading about a theory or being taught the theory in class, students should offer hypotheses for simple observations and create models or theories for more complex phenomena or relationships. There should be many questions in texts asking students to construct a hypothesis, model, or theory to account for something they have first observed. Then it is reasonable to conclude such material with the scientist's version of the theory or model.

The Design of a Reform in Science Education

The challenge then is to create a reform project that spaces learning over several years and moves from concrete experience to abstraction. Using a spiral approach the same concepts, principles, laws, and theories are studied at successively higher levels of abstraction, thus helping students to construct their own knowledge.

As an example, periodicity is a fundamental topic in science. In the project, students begin to look at periodic behavior through an example of the simple pendulum. At the 7th grade level, this includes observing that a pendulum takes about the same time to swing back and forth for a small swing as for a large swing; no measurements are made. Students then observe that when a lead bob is swung simultaneously with a wooden bob, using pendula of the same length, both pendula move together; one does not swing more quickly than the other. From this evidence, students can conclude that the time for a swing (not yet named a period) is independent of the bob's mass. They then try pendula of different lengths and observe that a longer pendulum takes more time to swing. These qualitative exercises provide experiences with the phenomenon. A qualitative relationship has been established.

At the 8th grade level, students again look at pendulums. This time they use a stopwatch to time how long the bob takes to move over and back once. They now give a name to this time, calling it the period of the pendulum. They repeat many of the same experiments that they did at grade 7, but this time they make measurements. They learn the meaning of amplitude, pendulum length, and bob mass. They plot a graph of period versus length, and from that graph *they predict* the period of the pendulum having a given length. Thus, they create a summary of an empirical law, in the form of a graph, and they make a prediction from that law and test the prediction.

At the 9th grade level, students conduct the experiment again, but this time they try different functions of the length, until they get a straight-line graph. They find this as a square root of the length. Thus, they conclude with an equation,

$$T = 0.20\ L$$

This empirical law is more specific than the empirical law determined as a graph in grade 8.

Finally, in the 11th or 12th grade, students use Newton's Second Law of Motion to derive the equation for the period of a pendulum, using a theory to account for what was observed empirically at earlier levels. The new law is more general. Unlike the previous empirical laws, which were valid only on Earth, the new equation,

$$T = 2\ L/g$$

allows us to find the period of a pendulum on the moon or somewhere else where the acceleration due to gravity is different.

This spiralling sequence on the pendulum occurs in the context of many other topics being learned. There are applications of the concepts and extensions to other areas, such as the mass on a spring, and vibrations in atoms and molecules, leading to heat and temperature as understood by kinetic theory.

Since students construct their own knowledge, they must start from their own preconceptions. Those preconceptions are exhibited mainly as metaphors or analogies young people have acquired to account for what happens in their world. In the pendulum experiments, students' preconceptions are such that they expect pendula having different mass bobs to swing with different periods. Students therefore face evidence from nature that their preconceptions must change.

Preconceptions, Metaphors, and Analogies

Student metaphors and analogies are often quite successful in accounting for the kinds of natural phenomena they encounter. The metaphors or analogies therefore cannot properly be called misconceptions or be given other pejorative descriptions.

Different cultures create quite different metaphors to account for nature—sometimes due to language differences, sometimes due to culture. For example, Asian languages are highly symbolic. It is probably no accident that Asians are particularly good at scientific theory where symbols are dominant. Some cultures, such as that of African-Americans, are rich in verbal metaphors. One must learn to listen carefully to understand the metaphors and how well they do account for some phenomena.

The research shows that preconceptions often are correct, but the quantities described have been transposed (Driver, 1984-1989). For example, many preconceptions about motion are due to the use of force when the person means energy.

Student Motivation

A fundamental requirement for an effective science education reform effort is to provide the necessary student motivation. Motivation is strongly connected to the degree to which provisions for other aspects of the research on learning have been accommodated. Hands-on experience with phenomena before terms are defined or concepts named is motivating. Sequencing over time at successively higher levels of abstraction provides a familiar experience which is motivating. Challenging student preconceptions by providing experiences for which their metaphors must be adapted (in the direction of those used by scientists) is also motivating.

Supporters of the Science/Technology/Society (STS) approach (Harms and Yager, 1981) stress relevance as a motivating factor; they believe that students must construct their knowledge of science by formulating and testing solutions to problems that occur in their daily lives, or in society. It is assumed that this will force consideration of, and motivate learning for, basic concepts, principles, and laws of science, which are then applied immediately to the problem or issue at hand. Relevance may well be a key component of good motivation, but practical problems are often very complex, and variables identified are almost impossible to isolate and control. Student interest in personal or societal problems is highly individual, and group learning in a classroom setting appears very difficult. The range of problems, issues, and concerns could easily spread into areas beyond the natural sciences, leading to a blurring of distinctions in areas where such distinctions are very important, such as between science and technology, or between science and philosophy and religion. Even though the author is convinced for many other reasons that the STS approach is flawed, it nevertheless needs a fair test. For that reason two of our project centers use an STS approach.

A Reform Project Begins

An article in the January/February 1988 issue of *NSTA Reports!*, which set down the basic project idea, created intense interest. The NSF, also showing great interest, issued a small grant for NSTA to organize and plan a large reform project. A first planning meeting was attended by some 110 scientists and science educators; participants included a Nobel Laureate in physics, association representatives in math, technology, and other subject-matter areas, teacher educators, science teachers, science educators, and even a Vice-President of the Academy of Sciences of the USSR.

At the meeting, the basic idea of spacing out the curriculum and sequencing from concrete to abstract was vastly modified, as if the basic idea were like a Christmas tree to which each participant added a few ornaments. The author could hardly recognize the original design when they were through. However, the model that emerged received general approval probably because participants could all see ways of fitting their own notions of reform into the structure. Thus began the SS&C project.

The SS&C project is now funded by both the Department of Education and the National Science Foundation. Some $13,000,000 in federal money has been invested in the project, along with considerable state and local resources. The project is underway in California, Texas, Iowa, North Carolina, Alaska, and Puerto Rico. Centers are attempting to provide science in four subjects every year—physics, chemistry, biology, and earth and space science—and to every student in grades 7-12. The middle level grades use integrated curricula in some places, STS curricula in others, and the teaching of separate subjects in others. At the high school level the subjects will be learned as separate disciplines, but with strong coordination. Each center is trying to build into their curricular design and learning materials the results of research described previously. The entire project is being documented and evaluated by Iris Weiss at Horizon Research, Inc., Chapel Hill, North Carolina, and the NSTA is serving as a national coordination center for the project. Because of the project's complexity, Arnold Strassenburg of SUNY at Stony Brook, New York, has accepted responsibility to direct the project.

Student Performance-Based Assessment and Compact-Disc Interactive

An interesting component of this project, and one in which the author has retained a continued interest, is student assessment. NSTA has a $400,000 grant from the Department of Education to produce a prototype CD-I disc designed to administer performance-based assessment.

Although CD-I hardware has just been released in the United States, the medium is ideal for student assessment. It has been sold in retail stores in the United States since October 1991. CD-I was chosen for the prototype because it will be a widely available and inexpensive consumer product.

The prototype disc guides students as they carry out tasks with real objects and phenomena. Students work alone, and the system tracks their responses, so that patterns of preconceptions, strategies, and techniques for solving problems can be identified.

The system will offer tasks or items at several levels of complexity, so that the item is more like a game than a test. All students can succeed at the lowest level of the item; as the item gets progressively more difficult, more students are exited (temporarily, for they can come back later) from the system. The items are designed so that only remarkably talented students can respond successfully at the top level. Thus, in most cases, nobody fails the item and nobody fully masters it! Students are neither arbitrarily discouraged nor inappropriately convinced of their superiority.

Script Summary for a Sample Item

To illustrate this multilevel concept, let us review the item being developed for the concept of density. At the lowest level, Tier 1, this item requires only that the student recall and use the basic formula for density, given values of mass and volume. The student is given two objects, one metal, the other wood, each with tags showing mass and volume. This is typically the level that most existing tests use to assess knowledge of density. The CD-I system guides the student through the solution of the problem, offering suggestions and other help. It also records responses.

At the next level, Tier II, the student must find the mass of a metal object using a balance and the volume of the object by determining its displacement in water. The student then finds the density. This level requires somewhat more knowledge and skill than Tier 1. Again, the student is given real objects and apparatus, and the CD-I serves mainly to guide the activity and offer assistance, while it records responses.

At Tier III, the student is given a regular solid, a balance, and a metric ruler. The student must measure the object's dimensions, find its volume, measure its mass, and finally calculate its density. The CD-I guides the student through these tasks, recording the steps taken and decisions made.

At Tier IV, the meaning of the concept of density is emphasized more. The earlier tasks comprised largely measurement and computation. In this tier, the student confronts the problem of determining the mass of a huge rectangular block of granite. It is a block so large that it could not be put on a scale. The student is given a small, irregularly shaped piece chipped off the granite block and told to use it to find the mass of the large block, knowing only its dimensions. The student is also given a graduated cylinder, water, and a balance. The student may get help as he or she progresses through the solution, and the system tracks strategies and responses. The student must find the volume of the sample rock, measure its mass, find its density, and then use the fact that density is mass per unit volume, along with the known dimensions of the larger block, to find first the volume, and then the mass of the larger block. This is a nontrivial task, and assesses understanding at a level substantially above what is typically done in testing.

Tier V goes further toward abstraction and requires an understanding of the concept of density and of the quantities associated with density. The student is given a rectangular piece of aluminum foil, a microbalance, and a meter stick. The student is told the density of the aluminum and asked for the thickness of the foil. The student must find the mass of the foil and its surface area (from length times width, which he or she must measure), and then must think through the necessary relationships, including that of volume as a function of mass and density and volume as a function of surface area and thickness.

Tier VI, the most difficult level, requires the student to determine the gold content of a piece of jewelry. The student is given a gold ring or pendant, a microbalance, a graduated cylinder with water, and the densities of gold and copper, of which the jewelry was alloyed. The task is to find the percentage of gold in the ring or pendant. This requires that the student find the volume of the object, weighing it in and out of water, to find the mass of water displaced, and therefore the volume of water displaced, and solve two simple simultaneous equations which the student must set up independently.

The CD-I program guides and helps students through the tiers; help becomes less extensive at the higher tiers. Students are exited from the system when their demands for assistance exceed some predetermined amount.

The system can simulate the performance students do with real objects, and motion can be slowed or calculations made or graphs prepared with the system. Thus, there is a dual set of things happening: student performance and system simulation or analysis.

Parallel Assessment of Cognitive Understanding

In addition to functional knowledge (student performance), the system will test for cognitive knowledge. There will be parallel items, involving an unusual adaption of a branching of questions, designed to find out how well understood the concepts are. For density, a series of such questions will be asked, testing whether students can arrive at satisfactory answers when they do not have to actually perform anything in the lab. In this part of the assessment, preconceptions and strategies will again be tracked.

All of the test items will be designed to assess depth of understanding. The goal is to find out how students know something is true, what evidence they have for some belief, and how they would go about learning something new. Facts and information will be strongly subordinate to the questions that formed the basis of the curriculum framework: What do we mean? How do we know? and Why do we believe?

In analyzing assessment results, it will be very interesting to see how functional and cognitive knowledge correlate. It will also be interesting to look at differences in cultures to see patterns of metaphors and analogies used in identified preconceptions and strategies.

When the prototype CD-I disc is complete and has been tested, a series of science test items using CD-I will be developed for grades 7-12. The CD-I technology can repeat an enormous number of permutations of protocols, precisely and with infinite patience. This is an inexpensive and powerful method of administering performance-based assessment. If good tasks can be designed, this improved assessment will drive the curricula to include good science education. And that is what all of us want.

Conclusions

The SS&C project is a major reform initiative. It is broadly based throughout the United States, and interest in implementing features of SS&C increases daily. New grants have already been issued to expand the project, and many schools are adopting parts of the program even before its field tests have been completed. Student assessment, as well as instructional programming, will ultimately require extensive use of computers, CD-I, and other technologies.

The fundamental structure of SS&C, with its focus on greater depth of understanding and less coverage, with appropriate sequence and consideration for developmental levels of students and of their preconceptions, may infuse other projects and reform initiatives. Whether diffusion is accompanied by loss of meaning and significance will only be told by the future. But if we achieve even modest goals, such as having science educators recognize the difficulties of learning something with understanding, we will have achieved much. In this project, and in what we teach students, we should heed Conant's observation:

"The stumbling way in which even the ablest of the scientists of every generation have had to fight through thickets of erroneous observations, misleading generalizations, inadequate formulations, and unconscious prejudice is rarely appreciated by those who obtain their scientific knowledge from textbooks" —James Bryant Conant (1893-1978), *Science and Common Sense*, 1951.

References

Arons, A.B. (1983). Achieving wider scientific literacy. *Daedalus*, Spring 1983.

Arons, A.B. (1990). *A Guide to Introductory Physics Teaching*, John Wiley & Sons, New York.

Arons, A.B., and Karplus, R. (1976). Implications of accumulating data on levels of intellectual development. *American Journal of Physics 44*: 396.

Bruner, J.S. (1960). *The Process of Education*, Harvard University Press, Cambridge, Massachusetts.

Dempster, F. (1988). The spacing effect: A case history in the failure to apply the results of psychological research. *American Psychologist 43*: 627.

Driver, R. (1984-1989). *The Children's Learning in Science Project,, Monographs on Preconceptions in Science*. Center for Studies in Science and Mathematics Education, Department of Education, University of Leeds, Leeds, England.

Harms, N.C. and Yager, R.E. (Eds.) (1981). *What Research Says to the Science Teacher, Vol. 3*, NSTA, Washington, D.C.

Janke, B. (1983, 1987, 1988). *Science and Technology in School Curricula: Case Studies 1-4, Course of Study for Lower Secondary Schools in Japan*, UNESCO, Paris, France.

Karplus, R. (1977). Science teaching and the development of reasoning. *Journal of Research in Science Education 14*: 169.

Linn, M.C. (1986). In Dillon, R.F. and Sternberg, R.J. (Eds.), *Science Cognition and Instruction*, Academic Press, New York.

McDermott, L.C. (1984). Research on conceptual understanding in mechanics. *Physics Today 37*: 24.

Piaget, J. (1973). *To Understand is to Invent: The Future of Education*, Grossman Publishers, New York.

Resnick, L.B. (1987). *Education and Learning to Think*, National Academy Press, Washington, D.C.

Szekely, B.B. (Ed.) (1987). New soviet secondary school science curricula. *Soviet Education 29*, No. 5-6, March-April 1987, M.E. Sharpe, Armonk, New York.

Wirszup, I. (1986). Soviet secondary school mathematics and science programs. In Sinclair, C. (Ed.), *The Status of Soviet Civil Service: Proceedings of the Symposium on Soviet Scientific Research*, NATO Headquarters, Brussels, Belgium, September 24-26, 1986, Martinus Nijhoff Publishers, Dordrecht/Boston/Lancaster.

SS&C LEARNING STATEMENTS

BEST COPY AVAILABLE

SPACING CONTENT IMPROVES LEARNING

One component of the SS&C project is a commitment to appropriate spacing and sequencing of science content. Dempster's article reviews psychological studies which have examined the spacing of content. Psychological research has identified a phenomenon called the "spacing effect"—that is, "for a given amount of study time, spaced presentations yield substantially better learning than do massed presentations." Yet, as Dempster points out, this phenomena has not received widespread classroom application. He considers nine possible impediments to the implementation of the research findings and concludes that none of them justify the absence of implementation.

Although the studies that Dempster reviews most clearly address the presentation of content on the classroom level and especially to learning "vocabulary," "scientific terms and arithmetical rules," he suggests that recognition of the spacing effect would lead to comprehensive changes in curricula. Similar to Aldridge, Dempster observes that, in comparison to American educational practice, Soviet mathematics textbooks present material with a much more distributed method. Dempster responds to charges that the spacing effect may apply only to the rote learning and memorization tasks currently under intense criticism in the education community. He states that "memory is of central importance to any complex intellectual activity." In other words, Dempster implies, the spacing effect may aid students in integrating knowledge and in constructing abstract concepts and theories.

THE SPACING EFFECT: A CASE STUDY IN THE FAILURE TO APPLY THE RESULTS OF PSYCHOLOGICAL RESEARCH

Frank N. Dempster

The spacing effect—which refers to the finding that for a given amount of study time, spaced presentations yield substantially better learning than do massed presentations—is one of the most remarkable phenomena to emerge from laboratory research on learning. It is remarkable in several respects. First, the spacing effect is one of the most dependable and replicable phenomena in experimental psychology. Second, it is remarkably robust. In many cases, two spaced presentations are about twice as effective as two massed presentations (e.g., Hintzman, 1974; Melton, 1970), and the difference between them increases as the frequency of repetition increases (Underwood, 1970). Moreover, demonstrations of achievement following massed presentations often are only slightly higher than that following a single presentation (e.g., Melton, 1970). Third, the spacing effect is truly ubiquitous in scope. It has been observed in virtually every standard experimental learning paradigm, with all sorts of traditional research material (Dempster, 1987a; Hintzman, 1974; Melton, 1970).

With all of these characteristics in its favor, the spacing effect would seem to have considerable potential for improving classroom learning. However, there is little evidence that this potential has been realized. Neither American classrooms nor American textbooks appear to implement spaced reviews in any systematic way, and by comparison, Soviet mathematics textbooks provide a much more distributed method of presentation than do their American counterparts (Stigler, Fuson, Ham, & Kim, 1986). Nor is there much evidence that the next generation of educators is being better informed. In a recent sampling of practitioner-oriented textbooks suitable for use in teacher education programs, I found either little or no mention of the practical benefits of the spacing effect, and in some cases the spacing effect was confused with other phenomena (e.g., Good & Brophy, 1986; Mayer, 1987; Slavin, 1986; Woolfolk, 1987). One well-known educator, in fact, advised against spaced practice at least in the early stages of learning (Hunter, 1983).

---

*This article first appeared in **American Psychologist** (1988), Volume 43, Number 8, pp. 627-634. Copyright © 1988 by the American Psychological Association. Reprinted by permission.*

Why is it that research findings that appear to have significant implications, such as the spacing effect, often are not utilized by teachers and curriculum makers? In general, the problem is that there is no well-developed implementation model, nor is there a standard methodology for analyzing the conditions that foster the transfer of knowledge from the laboratory to the classroom (see Hosford, 1984, for a discussion). Obviously, issues regarding the utilization of findings from basic research are complicated, and there are many potential impediments to the implementation of research findings in the classroom. In this article, I explore nine potential impediments, all of which seem reasonable at first glance, in an effort to determine which, if any, apply to the spacing effect.

Impediments to Application

The Phenomena Has Not Been Known Long Enough

Although the time lag between discovery and application varies greatly, some considerable period of time often intervenes between the publication of research findings and their application. In the case of the spacing effect, however, a considerable period of time already has passed since its initial documentation. The spacing effect was known as early as 1885 when Ebbinghaus published the results of his seminal experimental work on memory. With himself as the subject, Ebbinghaus found that for a single 12-syllable series, 68 immediately successive repetitions had the effect of making possible an errorless recital after seven additional repetitions on the following day. However, the same effect was achieved by only 38 distributed repetitions spread over three days. On the basis of this and other related findings, Ebbinghaus concluded that "with any considerable number of repetitions a suitable distribution of them over a space of time is decidedly more advantageous than the massing of them at a single time" (Ebbinghaus, 1885/1913, p.89). Jost, also working with nonsense syllables, reported similar findings and in 1897 formulated what was to become known as Jost's Law: "If two associations are of equal strength but of different age, a new repetition has a greater value for the older one" (McGeoch, 1943, p.140).

In 1928, Ruch published a review of dozens of studies of the spacing effect. Although interpretation of the results of these studies (e.g., Dearborn, 1910; Perkins, 1914; Pyle, 1913; Starch, 1912) is complicated by other, potentially confounded variables, the results tend, in general, to confirm the earlier work by Ebbinghaus and by Jost. Thus, published reports of the spacing effect have been in existence since the latter part of the 19th century and the early part of the 20th century.

The Phenomenon Has Not Received Recent Documentation

In the absence of recent documentation, research findings may seem stale or anachronistic, but, as most, if not all, students of the learning literature know, the spacing effect has been well-documented in recent times. Many studies of this phenomenon were published during the 1960s and the 1970s, as reviews by Hintzman (1974), Melton (1970), and Glenberg (1979) attest.

Although much of the research included in these reviews was reminiscent of the work of Ebbinghaus in using easily analyzable simple verbal units, the fruits of this research are considerable from any perspective. For example, the ubiquitous, highly replicable character of the spacing effect fostered the notion that its existence must be telling us something important about memory (e.g., Hintzman, 1974). Also, it clearly demonstrated that the Total Time Law, which states that the amount learned is a direct function of study time regardless of how that time is distributed, was in deep trouble or at least in need of a major overhaul (Melton, 1970; Underwood, 1970).

More recently, the spacing of repetitions has been the subject of studies reported in a variety of journals, including some with an applied perspective (Bahrick & Phelps, 1987; Cuddy & Jacoby, 1982; Dellarosa & Bourne, 1985; Dempster, 1987b; Elmes, Dye, & Herdelin, 1983; Glenberg & Lehmann, 1980; Glover & Corkill, 1987; Toppino & DiGeorge, 1984; Toppino & Gracen, 1985). Thus, documented evidence of the spacing effect has appeared in the literature continually for the past 100 years.

The Phenomenon Cannot Be Linked to Issues of Current Concern to Educators

I agree with Glaser (1982) that research knowledge is most likely to inform educational practice if it can be related explicitly to large-scale educational issues or macro variables. Whereas the relationships between the rather fine-grained analyses of learning and memory conducted by researchers and the practice of education usually are not self-evident to the teacher and must be illuminated by the psychologist, this task should be relatively straightforward in the case of the spacing effect. After all, the spacing effect has immediate and obvious implications for how time in the classroom may be distributed optimally. In the wake of recent critiques and studies of schooling, such as *A Nation at Risk* (National Commission on Excellence in Education, 1983), *Time to Learn* (Denham & Lieberman, 1980), and *Perspectives on Instructional Time* (Fisher & Berliner, 1985), the use of time in the classroom has become a major educational concern.

The Phenomenon Has Not Been Demonstrated Satisfactorily in School-Like Activities

Such demonstrations are arguably the most important bridge between basic research and educational practice. In the case of the spacing effect, however, this bridge seems to have been crossed. Several demonstrations of the spacing effect reviewed by Ruch (1928) were, as he put it, "intended for schoolroom application" (p. 20). One of the most interesting of these effects, from an educational perspective, is that of Pyle (1913), who had a group of third-graders drilled in addition, either twice a day for 5 days (once in the morning and once in the afternoon) or once a day for 10 days. Their improvement in recall of addition facts, which was decidedly in favor of the latter instructional method, provided perhaps the earliest experimental confirmation of William James' (1901) advice to teachers and students that it is better to repeat an association on many different days than again and again on just a few days (p. 129).

Most of the early demonstrations of the spacing effect "intended for schoolroom application" focused on text processing tasks, and in three recent studies text processing again has been the focus. In a study by Kraft and Jenkins (1981), subjects attempted to free-recall the title and one idea unit from each of a series of twice-presented passages, with each repetition separated by lags of up to eight intervening passages. As in much standard verbal learning research, recall was a linear and much improved function of lag. One practical limitation of this study, of course, is that students rarely are asked to recall so little from a series of passages.

In the second of these recent spacing effect studies, Dempster (1986) contrasted lags—defined as the interval between two opportunities to read a passage of text material—of

30 seconds, 5 minutes, 20 minutes, and 48 hours in one experiment, and 5 minutes and 30 minutes in a second experiment. In the first experiment, subjects in the 48-hour condition recalled significantly more idea units than did subjects in either of the two shortest lag conditions. In the second experiment, recall was significantly higher in the 30-minute condition than in the 5-minute condition. In both experiments, the recall advantage associated with the best performing group was about the same (between 25% and 30%), even though different text passages were used in the two experiments. Thus, the effect was both robust and replicable. Finally, Glover and Corkill (1987) observed the spacing effect (0 lag versus a 30-minute lag) in subjects' memory for paragraphs they read as well as for brief lectures (125 words) they heard.

In addition to text recall, spacing effects have been demonstrated in programmed instruction, where the objective has been the learning of science and mathematical concepts. In one study, the meanings of a series of programmed scientific terms were learned much more effectively when repetitions were spaced than when they were massed (Reynolds & Glaser, 1964). In another study, arithmetical rules presented by a computer-assisted instruction system and expressed as verbal statements were learned better when reviews occurred one and seven days after original learning than when they occurred one and two days after original learning (Gay, 1973, Experiment 2).

Finally, spacing effects have been found in vocabulary learning. In a study by Dempster (1987b), 38 uncommon English words and their definitions were presented three times, either with each repetition of any given word separated by every other word (i.e., each repetition of a word was separated by 37 other words or 4 minutes, 19 seconds) or with each repetition of a word massed in succession. In addition, the words were presented either with or without sentence contexts. The results were quite clear. In three experiments in which spaced versus massed presentations were evaluated in this manner, spaced presentations yielded substantially higher levels of vocabulary learning than did massed presentations. In some cases, in fact, the number of word meanings recalled was over 50% greater under spaced conditions than under massed conditions.

In Dempster's (1987b) study the retention interval averaged less than an hour—short, from a practical perspective. Thus, the question might be asked, "Would spacing effects emerge in the retention of vocabulary words tested at much longer intervals?" Fortunately, there are data relevant to this question. Bahrick and Phelps (1987) tested 35 individuals who had learned and relearned 50 English-Spanish word pairs for recall and recognition after an interval of eight years. One variable of interest was the interval between successive relearning sessions—either 30 days, 1 day, or 0. The data show that the intersession interval had a very large effect on recall, with the recall probability associated with the 30-day interval about 2.5 times the probability associated with the zero interval. In turn, the 1-day interval was associated with much better retention than the zero interval. For both comparisons, the effect on recognition, exclusive of recall, was much less pronounced. Moreover, their data clearly indicate that even five or more presentations are unlikely to facilitate long-term retention if the interval between successive presentations is one day or less. With respect to the educational implications of their study, the authors concluded that long-term retention would almost certainly be enhanced if foreign language courses make certain that students independently retrieve target information at intervals that are as long as 1 month, over a period of several years, instead of the more typical intervals of 1 to 2 days over periods of from 10 to 15 weeks (p. 349).

There Are Serious Discontinuities in the Literature on the Spacing Effect

Another possible reason for the failure to apply the results of research on the spacing effect is that there are serious discontinuities in the literature on the spacing effect, such that most recent studies seem uninformed by the research of earlier ones. Although the spacing effect has a lengthy published history, there are discontinuities of this sort. For example, much of the important work from an applied perspective done in the early 1900s is not cited by studies published later. More generally, most recent studies tend to give the impression that, with the exception of the work of Ebbinghaus and Jost, all we know about the spacing effect dates only as far back as the 1960s. I cannot help but wonder how widespread a problem this is in the learning literature, and what sorts of consequences it has for the evolution of the science of learning and its application. Why is it that we occasionally—and perhaps frequently—give up on, or simply lose interest in, a phenomenon before we have definitive answers to basic questions and, then much later, return to the phenomenon as though we had just recently discovered it?

Upon reflection, this ahistorical character of research on the spacing effect would seem to have at least two unfortunate consequences, either of which could impede application. From a scientific perspective, such research is less likely to be as cumulative—where cumulative refers to empirical laws and theoretical structures building on one another so that later developments extend and unify earlier work (Hedges, 1987)—as would be more historical research. From a practical perspective, it will seem as if the spacing effect has not weathered continuous scrutiny over a lengthy period of time.

There are, of course, many possible reasons for discontinuities in the literature of a phenomenon. In the case of the spacing effect, however, three appear to stand out. First, for the most part, studies conducted from an applied perspective and those conducted from a basic research perspective constitute two distinct streams of research. For example, although widely cited reviews by Crowder (1976), Glenberg (1979), Hintzman (1974), and Melton (1970) report dozens of studies of the spacing effect using simple list learning materials, none of the many studies using more complex material with clearer classroom analogues are reported. This is the case despite the fact that some still viable accounts of the spacing effect (see for example, Cuddy & Jacoby, 1982, and Hintzman, 1974) were anticipated and supported in earlier studies done from a more applied perspective (see, for example, Ausubel, 1966).

Second, different terminology has been used to refer to similar, though distinguishable, phenomena—a situation that seems to have created some confusion. For example, a number of writers have distinguished between the "spacing effect" and both the "Melton" or "lag effect" and the "massed-versus-distributed practice" effect. Lag effects have been said to occur when performance improves as a function of the number of intervening items between successive presentations, whereas massed versus distributed practice effects often have been restricted to comparisons between spacings of zero (massed practice) and all spacings greater than zero (distributed practice). The use of this terminology, however, is somewhat uneven. Some researchers have used these terms interchangeably or have included other variables (e.g., length of period of study, retention interval) in their characterizations of the spacing effect (e.g., Ruch, 1928). Also, much of the massed versus distributed practice research, which has focused on perceptual motor skills tasks and lists of nonsense syllables, has yielded weak effects of spacing. According to Underwood (1961), who reviewed 10 years of distributed practice research, "Even under the most favorable conditions for facilitation by distributed practice, one could not recommend its use in an applied setting" (p. 230). Apparently following his lead, some older educational psychology texts advised that there were no clear practical implications to be drawn from distributed practice research (e.g., DeCecco, 1968; Mathis, Cotton, & Sechrest, 1970).

Third, the spacing effect is just one of a family of similar, though less thoroughly investigated, phenomena that are occasionally confused in the literature. One such phenomenon is the so-called "test-spacing" effect, which refers to the fact that spaced tests, particularly tests with intertest intervals of an expanding nature, result in greater retention than do massed testings (Landauer & Bjork, 1978; Rea & Modigliani, 1985; Spitzer, 1939). Another related phenomenon has been observed when once-presented written exercises or materials in a short course in statistics either are spread over the course of several sessions or are presented in a single session—a situation that is analogous to "cramming" for a test. In this case, students learn more when the material is distributed over several sessions (Bloom & Shuell, 1981; Smith & Rothkopf, 1984). Finally, Reder and Anderson (1982) found that, with total study time equated, repeated, well-spaced presentations of a text were more effective than was a single, longer presentation. Similarly, Edwards (1917) compared groups who studied various school materials, including history and geography, six and one-half minutes continuously or with the same amount of time divided into a study period of four minutes followed some days later by a review of two and one-half minutes. Lag and test intervals, difficulty of the material, and age of the subjects were so variable that it is difficult to interpret the results; however, without exception they favor the repeated, spaced-study groups.

To complicate matters further, research on allied phenomena also has an ahistorical character. For example, Rea and Modigliani (1985) failed to cite Spitzer (1939), even though Spitzer's work showed that if the interval between original learning and the first test in a series is too lengthy, test spacing effects are likely to be vitiated. Also, Reder and Anderson (1982) failed to cite Edwards (1917), and neither Bloom and Shuell (1981) nor Smith and Rothkopf (1984) seemed aware of the work of Ash (1950), who found only minor differences in retention between groups treated like theirs.

Too Many Studies Using School-Like Activities Have Failed to Show the Spacing Effect

Although the spacing effect is one of the most dependable phenomena in the learning of standard verbal learning lists, there have been more than a half-dozen documented failures to observe the spacing effect in tasks with classroom analogues. The results of these studies make it quite clear that the spacing effect is subject to certain not fully understood boundary conditions. Specifically, five sorts of boundary conditions are suggested. First, it has been found that under certain circumstances spaced presentations are no better than (Austin, 1921) and sometimes even worse than (Gordon, 1925) massed presentations in tests of immediate recall. For example, Austin found that massed readings (e.g., five times in one day) of text material proved as effective as spaced readings (e.g., daily for five days) in tests of immediate recall, whereas the spaced readings were much more effective in delayed tests, particularly if they came two to four weeks after learning. Second, it has been found that massed practice often is more efficient for certain simple, isolated skills, such as writing the products of number pairs as rapidly as possible (Thorndike, 1916). Third, evidence from traditional learning research suggests that the spacing effect may not apply to preschool age children, although it does emerge in a robust manner by age seven (Toppino & DiGeorge, 1984). Fourth, two studies have shown that the spacing effect can be eliminated if paraphrased rather than verbatim versions of the repeated materials are used (Dellarosa & Bourne, 1985; Glover & Corkill, 1987).

Finally, the results of a number of studies seem to suggest that beyond a certain lag interval, further increases in lag are not always associated with further increases in learning. For example, English, Wellborn, and Killian (1934) found that four readings of a text at three-hour intervals were associated with better learning than four consecutive unspaced readings; however, readings at three-hour intervals were no better than readings at either one- or three-day intervals. Similarly, Lyon (1914), Peterson, Ellis, Toohill, and Kloess (1935), and Sones and Stroud (1940) reported essentially no differences in retention between groups with rereading reviews spaced 1 and 7, 1 and 9, and 1 and 17 days after original learning. These findings were later corroborated by Ausubel (1966) and by Gay (1973).

Of these boundary conditions, the last two seem the most serious and most puzzling in view of the fact that the spacing effect has been found in paraphrased material (Rothkopf & Coke, 1966) and in view of long-lag effects obtained with traditional verbal learning material (Glenberg & Lehmann, 1980) and in vocabulary learning (Bahrick & Phelps, 1987). With respect to the latter, it may be that under certain lengthy lag conditions, the usual benefits of spaced repetitions do not obtain because the results of initial processing efforts have been forgotten (see Lyon, 1914, and Sones & Stroud, 1940, for earlier discussions of this hypothesis).

In any case, the spacing effect, especially in text processing, cannot be taken for granted. Although there have been relatively few documented failures to obtain the spacing effect, they certainly could have raised enough doubt about the dependability of the effect to discourage application.

The Phenomenon Has Not Been Demonstrated Satisfactorily in the Classroom

With very few exceptions (Dempster, 1986; Glover & Corkill, 1987; Pyle, 1913), even research using educationally relevant materials has been conducted in the laboratory. Moreover, the rather simple learning situations created in classroom studies of the spacing effect fail to approach the complexities facing curriculum developers and teachers. Curriculum developers and teachers have to concern themselves with the design of instruction conveyed in classes that often meet every day for a school term, under conditions in which much if not most of the content is organized in ways that imply systematic movement through learning hierarchies, curricular sequences, and so forth. Thus, it may be argued that it is not at all clear what specific implications demonstrations of the spacing effect in simple, isolated classroom situations have for curriculum designers and teachers faced with decisions about how much material to include in a course, how to sequence it, and how to optimally phase in new material and phase out old material.

Arguably, the relative lack of applied research in educational settings is, from an educational perspective, the most serious shortcoming of research on the spacing effect. There is no substitute for applied research, and the absence of at least several convincing demonstrations of the spacing effect in ongoing classroom situations under naturalistic conditions may well have been an impediment to application.

Too Little Is Known About Actual Classroom Practice to Justify Widespread Application of the Spacing Effect

To the best of my knowledge, nothing has been published concerning the proportion of time in the classroom teachers usually devote to review (i.e., any re-presentation or practice activity pertaining to a particular educational objective), and to what extent reviews are massed as opposed to spaced. Moreover, the general nature of classroom review activities apparently has not been characterized. For example, to what extent are reviews verbatim or paraphrased? Because several studies have failed to show spacing effects in text processing when paraphrased rather than verbatim versions of the repeated materials were used (Dellarosa & Bourne, 1985; Glover & Corkill, 1987), the answer to this question is of practical importance.

In short, due to significant gaps in our understanding of classroom practice, the application potential of the spacing effect cannot be estimated with any precision. Accordingly, it may be assumed wrongly that efforts to implement the spacing effect would result in little benefit.

The Phenomenon Is Not Sufficiently Understood

Phenomena that are not well understood are likely to invite skepticism among practitioners—particularly those who are familiar with any one of several instances in which the application of a poorly understood finding has had extremely unfavorable consequences (e.g., thalidomide). For example, educators, who often give the impression of having a low regard for fact memorization, might feel that the spacing effect would interfere with the operation of "more laudable, higher mental processes," because it is exactly such memorization to which spacing applies most clearly.

In fact, the theoretical picture surrounding the spacing effect is confused and uncertain, despite numerous attempts at clarification (e.g., Dellarosa & Bourne, 1985; Hintzman, 1974). A recent case in point is a published failure to replicate findings implicating a component-levels interpretation of the spacing effect in standard verbal learning tasks (Toppino & Gracen, 1985). The component-levels hypothesis is one of a group of hypotheses that attribute the spacing effect to increasing independence of encoding events with increasing intervals between repetitions.

With more naturalistic material, theoretical work has been slow to develop, and little has been done in the way of theoretically derived hypothesis testing. However, it is worth noting that Ausubel (1966), in a rarely cited study, offered the following explanation for the advantages of spaced review in meaningful learning:

> In the first place, after a longer retention interval, when more material is forgotten, the learner is more highly motivated to profit from the opportunity for review. He is less likely to regard this opportunity as unnecessary and superfluous and is hence more disposed to take good advantage of it in terms of effort, attention, and concentration (p. 197).

Apparently, Ausubel (1966) was on the right track when he stressed the role of attention and effort in the spacing effect. Recent evidence indicates that some sort of attentional account of the spacing effect is at least as viable as any other account and well worth pursuing (Dellarosa & Bourne, 1985; Dempster, 1986; Elmes, Dye, & Herdelin, 1983; Magliero, 1983). For example, Dempster (1986) had subjects respond to a questionnaire administered following a recall test of a twice-read passage, with the two readings spaced either 30 minutes apart or 5 minutes apart. The questionnaire consisted of 10 items, each followed by a 10-point rating scale, which was designed to elicit self-reports of various cognitive and affective states and processes during reading and testing. Included were questions concerning levels of attention, interest, anxiety, rehearsal, and changes of interpretation from one reading to the text. The results were quite clear. Significant group differences emerged on only two of the items, specifically one asking the subjects to indicate how "interested" they were during the second reading—an affective state—and one asking them to indicate how much "attention" they paid during the second reading—a cognitive process. In both cases, the average ratings of students in the spaced 30-minute condition (those who also did best on the recall test) were higher than those in the massed 5-minute condition. Moreover, a correlational analysis, applied to the scores of both groups combined, revealed a significant correlation between recall and only one of the questionnaire items—that is, the attention paid during the second reading. Those who reported having paid more attention tended to have learned more from the text.

However, why should spaced presentations be more interesting (see also Elmes et al., 1983) and receive more attention than massed presentations? Here again, Ausubel (1966) seemed to have anticipated more recent developments when he suggested that it has to do with the relative accessibility of previous encodings (in his words, the activity of "trying and failing to remember material," p. 197). If a student receives massed presentations, the information learned during earlier presentations should be relatively easy to retrieve from memory during subsequent presentations. Thus, subsequent presentations should be relatively redundant or familiar and thus relatively boring. By contrast, if a student receives two well-spaced presentations, the information learned during initial presentations should be relatively inaccessible during subsequent presentations, which should heighten interest level and the amount or quality of attention subsequent presentations receive. According to this account, then, massed presentations are relatively ineffective because they may not actually result in much repetitive processing (see also Cuddy & Jacoby, 1982; Dellarosa & Bourne, 1985; Greeno, 1970; Jacoby, 1978; Underwood, 1970).

One implication of this account is that anything that increases the likelihood that a repetition will receive full processing, such as events that make it difficult to retrieve the results of prior encodings, should improve learning. Thus, this account helps to explain failures to obtain the spacing effect with paraphrased repetitions, that is, repetitions having a changed surface structure (Dellarosa & Bourne, 1985; Glover & Corkill, 1987), and under lengthy lag conditions (Ausubel, 1966; English et al., 1934; Gay, 1973; Lyon, 1914; Peterson et al., 1935; Sones & Stroud, 1940).

In sum, although recent studies have yielded some promising clues to the mechanisms underlying the spacing effect, our theoretical ignorance may have been and may continue to be an impediment to application or might contribute to inappropriate applications.

Summary and Conclusions

The spacing effect would seem to have considerable potential for improving classroom learning, yet there is no evidence of widespread application. In this article, I have considered nine possible impediments to the implementation of research findings in the classroom. Of the nine, five appear to apply to the spacing effect. These include the ahistorical character of research on the spacing effect, some failures to obtain the effect with school-like activities, a paucity of impressive classroom demonstrations of the phenomenon, limited knowledge of classroom practice, and an incomplete understanding of the psychological bases of the spacing effect. By contrast, the fact that the phenomenon (a) has been known for a long period of time, (b) has received recent documentation, (c) can be linked to current educational issues, and (d) has been shown to extrapolate to school-like activities suggests that the first four dimensions of analysis considered do not apply.

The following question now arises: To what extent *should* the list of plausible impediments to the application of the spacing effect in the classroom discourage application? After addressing each of these concerns, I conclude that we do know enough about the effect of spacing to make a very strong argument for application without any additional knowledge about the spacing effect or classroom practice.

Consider first the ahistorical and somewhat confusing character of research on the spacing effect. Given the long history of research on the spacing effect and the considerable recent documentation, the discontinuities in the literature on the spacing effect do not seem terribly important. In fact, the spacing effect has weathered continuous scrutiny over a lengthy period of time. Of course, the fact that the spacing effect is just one of a family of similar phenomena would seem only to strengthen and broaden the appeal of "spacing" in the classroom.

A second plausible reason for the absence of widespread application of the spacing effect is that there have been some failures to obtain the effect in school-like tasks. The spacing effect, in fact, does appear to be subject to certain, not fully understood boundary conditions. However, it would be unrealistic to expect the spacing effect to apply in every situation, and the relatively few failures to obtain the spacing effect seem trivial in light of its many demonstrations. For this reason and in the absence of any serious contraindications to the application of the spacing effect, the fact that the spacing effect does not always work hardly seems to justify resistance to its application.

Arguably, the most serious of the plausible impediments to the application of the spacing effect is the paucity of impressive classroom demonstrations of the phenomenon. Clearly, programmatic research on the effects of spacing in education settings is long overdue, as the results of such efforts would likely aid in its application. The most useful studies of this sort would be those involving curriculum design and classroom teaching that help shed light on the implications of the spacing effect for specific applied issues, such as homework, frequency of testing and feedback, learning hierarchies, mastery learning, and questioning strategies.

Although additional classroom research is desirable and even necessary in order to make the most effective use of spacing, I do not think we have to withhold all judgement on the applicability of the spacing effect at this point. For example, we know that children study vocabulary in a variety of subjects and that the spacing effect facilitates vocabulary learning. Likewise, children are exposed to numerous scientific terms and arithmetical rules while in school—each of which is a domain in which spacing has been found to be effective. When coupled with what we already know about the effects of spacing in text processing, there is ample reason to believe that the spacing effect will improve classroom learning in a wide variety of subject areas.

Although we do not know much about current classroom applications of the spacing effect, there are many reasons to believe that the spacing effect is under-utilized in the classroom in terms of its potential for improving learning. In addition to the reasons indicated earlier, there is the fact that the spacing effect is somewhat counterintuitive. The spacing effect typically refers to a phenomenon that occurs under conditions in which the retention interval between the last presentation and the test is held constant. Thus, one might reason that because the retention interval between the first presentation and the test is shorter under massed conditions, this condition should result in superior performance. Even experienced educators, when judging the instructional effectiveness of text passages, tend to rate prose in which the repetition of a given unit of information is massed as better than those in which it is spaced (Rothkopf, 1963).

In short, the spacing effect is neither intuitively obvious, nor well known among educators. Accordingly, it is reasonable to assume that those who become teachers, administrators, curriculum developers, or writers of reading series are ignorant of the spacing effect, just as many psychologists are not clear about the totality of educational situations that call for its application. Thus, our ignorance of actual classroom practice should not be interpreted to mean that widespread implementation of the spacing effect has little or no potential for improving classroom learning.

The final plausible impediment to the application of the spacing effect is that it is not well understood theoretically. However, in view of the absence of evidence linking the spacing effect to some undesirable psychological process or outcome, this impediment is extremely hard to justify. For example, there is no evidence that the spacing effect impairs the ability to conceptualize or to think critically. Even if it is found that the psychological basis of the spacing effect applies only or largely to memorization tasks, memory is of central importance to any complex intellectual activity.

Nevertheless, a fuller understanding of the spacing effect could eventually aid in its application and might help to avoid some inappropriate applications. Hence, I believe the search for the underlying cause(s) of the spacing effect should proceed. As part of this effort, some attempt should be made to determine if spaced presentations tend simply to increase the amount of information learned (the quantitative hypothesis), or if they tend to increase the learning of only certain kinds of information (the qualitative hypothesis). So far, these hypotheses have been examined only in relation to the more general issue of how repetitions improve learning, with mixed results (Annis & Annis, 1987; Bromage & Mayer, 1986: Mayer, 1983). They have not been examined in the context of comparisons of massed and spaced repetitions, although such comparisons might be fruitful from both a theoretical and a practical perspective.

Recently, I was told that the spacing effect has been "studied to death," and on another occasion that "we know all that we need to know about the spacing effect." Clearly, the spacing effect is one of the most studied phenomena in the 100-year history of learning research, and we *do* know a lot about it—enough to recommend application. However, it would be a mistake to do what these comments imply—which is simply to stop investigating the phenomenon. Although it may take some clever research to avoid diminishing returns, continued experimental study of the spacing effect can yield valuable information regarding its parameters and cause(s). Then too, applied research and widespread application should produce the data base needed to evaluate the particular conditions under which the spacing effect works best.

Author's Note: Preparation of this article was supported, in part, by a grant from the Barrick Research Council, University of Nevada, Las Vegas. I thank my colleagues John A. Bates and Doris Carey for their assistance, as well as Arthur Glenberg, Douglas Hintzman, and an anonymous reviewer for their helpful reactions to an earlier version of this article.

References

Annis, L.F. & Annis, D.B. (1987, April). *Does practice make perfect? The effects of repetition on student learning.* Presented at the annual meeting of the American Educational Research Association, Washington, DC.

Ash, P. (1950). The relative effectiveness of massed versus spaced film presentations. *Journal of Educational Psychology, 41,* 19-30.

Austin, S.D.M. (1921). A study in logical memory. *American Journal of Psychology, 32,* 370-403.

Ausubel, D.P. (1966). Early versus delayed review in meaningful learning. *Psychology in the Schools, 3,* 195-198.

Bahrick, H.P., & Phelps, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 344-349.

Bloom, K.F., & Shuell, T.J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research, 74,* 245-248.

Bromage, B.K. & Mayer, R.E. (1986). Quantitative and qualitative effects of repetition on learning from technical text. *Journal of Educational Psychology, 78,* 271-278.

Crowder, R.G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Cuddy, L.J., & Jacoby, L.L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning and Verbal Behavior, 21,* 451-467.

Dearborn, W.F. (1910). Experiments in learning. *Journal of Educational Psychology.* Englewood Cliffs, NJ: Prentice-Hall.

DeCecco, J.P. (1968). *The psychology of learning and instruction: Educational psychology.* Englewood Cliffs, NJ: Prentice-Hall.

Dellarosa, D., & Bourne, L.E. (1985). Surface form and the spacing effect. *Memory & Cognition, 13,* 529-537.

Dempster, F.N. (1986). *Spacing effects in text recall: An extrapolation from the laboratory to the classroom.* Manuscript submitted for publication.

Dempster, F.N. (1987a). Time and the production of classroom learning: Discerning implications from basic research. *Educational Psychologist, 22,* 1-21.

Dempster, F.N. (1987b). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology, 79,* 162-170.

Denham, C., & Lieberman, A. (1980). *Time to learn.* Washington, DC: National Institute of Education.

Ebbinghaus, H. (1913). *Memory* (H.A. Ruger & C.E. Bussenius, Trans.). New York: Teachers College. (Original work published 1885; paperback ed., New York, 1964).

Edwards, A.S. (1917). The distribution of time in learning small amounts of material. In *Studies in*

Psychology: Titchener commemorative volume (pp. 209-213). Worchester, MA: Wilson.

Elmes, D.G., Dye, C.J., & Herdelin, N.J. (1983). What is the role of affect in the spacing effect? *Memory & Cognition, 11*, 144-151.

English, H.B., Welborn, E.L., & Killian, C.D. (1934). Studies in substance memorization. *Journal of General Psychology, 11*, 233-260.

Fisher, C.W., & Berliner, D.C. (1985). *Perspectives on instructional time*. New York: Longman.

Gay, L.R. (1973). Temporal position of reviews and its effect on the retention of mathematical rules. *Journal of Educational Psychology, 64*, 171-182.

Glaser, R. (1982). Instructional psychology: Past, present, and future. *American Psychologist, 37*, 292-305,

Glenberg, A.M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7*, 95-112.

Glenberg, A.M., & Lehmann, T.S. (1980). Spacing repetitions over 1 week. *Memory & Cognition, 8*, 528-538.

Glover, J.A., & Corkill, A.J. (1987). Influence of paraphrased repetitions on the spacing effect. *Journal of Educational Psychology, 79*, 198-199.

Good, T.L., & Brophy, J.E. (1986). *Educational psychology* (3rd ed.). New York: Longman.

Gordon, K. (1925). Class results with spaced and unspaced memorizing. *Journal of Experimental Psychology, 8*, 337-343.

Greeno, J.G. (1970). Conservation of information-processing capacity in paired-associate memorizing. *Journal of Verbal Learning and Verbal Behavior, 9*, 581-586.

Hedges, L.V. (1987). How hard is hard science, how soft is soft science? *American Psychologist, 42*, 443-455.

Hintzman, D.L. (1974). Theoretical implications of the spacing effect. In R.L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 77-99). Potomac, MD: Erlbaum.

Hosford, P.L. (Ed.). (1984). *Using what we know about teaching*. Alexandria, Va: Association for Supervision and Curriculum Development.

Hunter, M. (1983). *Increasing your teaching effectiveness*. Belmont, CA: The Learning Institute.

Jacoby, L.L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior, 17*, 649-667.

James, W. (1901). *Talks to teachers on psychology: And to students on some of life's ideals*. New York: Holt.

Kraft, R.N. and Jenkins, J.J. (1981). The lag effect with aurally presented passages. *Bulletin of the Psychonomic Society, 17*, 132-4.

Landauer, T.K., & Bjork, R.A. (1978). Optimum rehearsal patterns and name learning. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical aspects of memory* (pp. 625-632). New York: Academic Press.

Lyon, D.O. (1914). The relation of length of material to time taken for learning and the optimum distribution of time. *Journal of Educational Psychology, 5*, 1-9, 85-91, 155-163.

Magliero, A. (1983). Pupil dilations following pairs of identical and related-to-be-remembered words. *Memory & Cognition, 11*, 609-615.

Mathis, B.C., Cotton, J.W., & Sechrest, L. (1970). *Psychological foundations of education*. New York: Academic Press.

Mayer, R.E. (1983). Can you repeat that? Qualitative effects of repetition and advanced organizers on learning from science prose. *Journal of Educational Psychology, 75*, 40-49.

Mayer, R.E. (1987). *Educational psychology: a cognitive approach*. Boston: Little, Brown.

McGeoch, J.A. (1943). *The psychology of human learning*. New York: Longmans Green.

Melton, A.W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal* *of Verbal Learning and Verbal Behavior, 9,* 596-606.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

Perkins, N.L. (1914). The value of distributed repetitions in rote learning. *British Journal of Psychology, 7,* 253-261.

Peterson, H.A., Ellis, M., Toohill, N., & Kloess, P. (1935). Some measurements of the effects of reviews. *Journal of Educational Psychology, 26,* 65-72.

Pyle, W.H. (1913). Economical learning. *Journal of Educational Psychology, 3,* 148-158.

Rea, C.P., & Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning, 4,* 11-18.

Reder, L.M., & Anderson, J.R. (1982). Effects of spacing and embellishment on memory for the main points of a text. *Memory & Cognition, 10,* 97-102.

Reynolds, J.H., & Glaser, R. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology, 55,* 297-308.

Rothkopf, E.Z. (1963). Some observations on predicting instructional effectiveness by simple inspection. *Journal of Programmed Instruction, 3,* 19-20.

Rothkopf, E.Z., & Coke, E.V. (1966). Variations in phrasing and repetition interval and the recall of sentence materials. *Journal of Verbal Learning and Verbal Behavior, 5,* 86-91.

Ruch, T.C. (1928). Factors influencing the relative economy of massed and distributed practice in learning. *Psychological Review, 35,* 19-45.

Slavin, R.E. (1986). *Educational Psychology: Theory into practice.* Englewood Cliffs, NJ: Prentice-Hall.

Smith, S.M., & Rothkopf, E.Z. (1984). Contextual enrichment and the distribution of practice in the classroom. *Cognition and Instruction, 1,* 341-358.

Sones, A.M., & Stroud, J.B. (1940). Review with special reference to temporal position. *Journal of Educational Psychology, 31,* 665-676.

Spitzer, H.F. (1939). Studies in retention. *Journal of Educational Psychology, 30,* 641-656.

Starch, D. (1912). Periods of work in learning. *Journal of Educational Psychology, 3,* 209-213.

Stigler, J.W., Fuson, K.C., Ham, M., & Kim, M.S. (1986). An analysis of addition and subtraction word problems in American and Soviet elementary mathematics textbooks. *Cognition and Instruction, 3,* 153-171.

Thorndike, E.L. (1916). Notes on practice, improvability and the curve of work. *American Journal of Psychology, 27,* 550-565.

Toppino, T.C., & DiGeorge, W. (1984). The spacing effect in free recall emerges with development. *Memory & Cognition, 12,* 118-122.

Toppino, T.C., & Gracen, T.F. (1985). The lag effect and differential organization theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 185-191.

Underwood, B.J. (1961). Ten years of massed practice on distributed practice. *Psychological Review, 68,* 229-247.

Underwood, B.J. (1970). A breakdown of the total-time law in free-recall learning. *Journal of Verbal Learning and Verbal Behavior, 9,* 573-580.

Woolfolk, A.E. (1987). *Educational psychology.* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

LEARNING IS NOT IMPROVED BY ABILITY GROUPING

Clearly, a key motivation of Bill Aldridge's first article was a frustration with the "systematic filtering" of able students out of the sciences. He observes that "the present sequencing of courses and tracking of students reinforce the accumulation of advantage for some students and effectively preclude others from later selecting science as a career." Aldridge adamantly argues that all students can and should study science. Both articles in this section address the issue of tracking---the grouping of students by ability---and how such grouping affects student achievement. Slavin reviews research on the effects of ability grouping on the achievement of secondary students. He discusses the most common form of ability grouping---between-class grouping---as well as other forms. It is important to note that both Slavin and Aldridge believe that patterns of course taking significantly affect achievement (i.e., students take chemistry after biology and physics after chemistry and that only certain students follow through the progression), but little research has been conducted in this area. Slavin's key conclusion is that "comprehensive between-class ability grouping plans have little or no effect on the achievement of secondary students." In other words, the common practice of grouping high-, average-, and low-achieving students with their respective counterparts is ineffective and, as Slavin points out, the practice is essentially "anti-democratic," "antiegalitarian," and not responsive to the instructional needs of a heterogeneous student body.

The second article, written by Manning and Lucking, takes up this issue of the instructional needs of the heterogeneous student body. The authors discuss cooperative learning---an alternative to traditional grouping plans and instructional systems. Manning and Lucking define cooperative learning as "techniques in which students work in heterogeneous groups of four or six members and earn recognition, rewards, and sometimes grades based on the academic performance of their groups." They discuss evidence indicating that cooperative learning can contribute positively to students' academic achievement, social skills, and self-esteem. For the authors, a particularly notable aspect of cooperative learning programs is their ability to improve relations within and between multicultural/multiethnic students. The research studies bear out the notion that "people who help one another by joining forces to achieve a common goal generally feel more positively about each other and are willing to interact more positively when performing collective tasks." Manning and Lucking's review also serves as an appropriate introduction to the subsequent section in this volume on encouraging equitable classrooms.

ACHIEVEMENT EFFECTS OF ABILITY GROUPING IN SECONDARY SCHOOLS: A BEST-EVIDENCE SYNTHESIS

Robert E. Slavin

For more than seventy years, ability grouping has been one of the most controversial issues in education. Its effects, particularly on student achievement, have been extensively studied over that time period, and many reviews of the literature have been written. In recent years, a comprehensive review of the achievement effects of ability grouping in elementary schools has been published by Slavin (1987), but only brief meta-analyses by Kulik and Kulik (1982, 1987) have reviewed the evidence on ability grouping and heterogeneous placement in secondary schools.

The purpose of this paper is to present a comprehensive review of all research published in English that has evaluated the effects of ability grouping on student achievement in secondary schools. *Secondary schools* are defined here as middle, junior, or senior high schools in the United States, or similarly configured secondary schools in other countries. Secondary schools can include grades as low as five, but they usually begin with sixth or seventh grades. Ability grouping is defined as any school or classroom organization plan that is intended to reduce the heterogeneity of instructional groups; in between-class ability grouping the heterogeneity of each class for a given subject is reduced, and in within-class ability grouping the heterogeneity of groups within the class (e.g., reading groups) is reduced.

Unlike the situation in elementary schools, the type of ability grouping used in secondary schools is overwhelmingly between-class grouping (McPartland, Coldiron, & Braddock, 1987). Several closely related forms of ability grouping are used. Sometimes students are assigned on the basis of some combination of composite achievement, IQ, and teacher judgments to a track, within which all courses are taken. For example, senior high school students are often assigned to academic, general, and vocational tracks; middle/junior high school students are often assigned to advanced, basic, and remedial tracks (in either case, the number of tracks and the names used to

---

This article first appeared in the ***Review of Educational Research*** *(Fall 1990), Volume 60, Number 3, pp. 471-499.*

Copyright © 1990 by the American Educational Research Association. Reprinted by permission.

describe them vary widely). This type of grouping plan is generally called tracking in the United States or streaming in Europe. It is an example of what Slavin (1987) called "ability-grouped class assignment." In addition to assignment to higher and lower sections of the same courses, tracking in senior high schools usually involves different courses or course requirements. For example, a student in the academic track may be required to take more years of mathematics than a student in the general track, or may take French III rather than metal shop.

A particular form of tracking often seen in middle/junior high schools is block scheduling, where students spend all or most of the day with one homogeneous group of students. Some schools rank-order students from top to bottom and assign them to, say, 7-1, 7-2, 7-3, and so on. Many senior high schools allow students to choose their track or to choose the level they wish to take in each subject, but in plans of this kind counselors tend to steer students into the level of classes to which they would have been assigned if the school were not allowing students a choice (Rosenbaum, 1978).

Another form of ability grouping common in secondary schools involves assigning students to ability-grouped classes for all academic subjects, but allowing for the possibility that students will be placed in a high-ranking group for one subject and a low-ranking group for another. In practice, scheduling constraints often make this type of grouping similar to plans in which all courses are taken within the same track. In some cases schools ability group for some subjects and not for others; for example, students may be in ability-grouped math and English classes but in heterogeneous social studies and science classes. Ability grouping usually involves higher and lower sections of the same course, but sometimes consists of assignment to completely different courses, as when ninth graders are assigned either to Algebra I or to general math. When high achievers are assigned to markedly different courses usually offered to older students (as when seventh graders take algebra), this is called acceleration. More commonly, high achievers may be assigned to "honors" or "advanced placement" sections of a given course, and low achievers may be assigned to special "remedial" sections.

Although between-class ability grouping is by far the most common type of ability grouping in secondary schools, forms of within-class grouping are also occasionally seen. These are plans in which students are assigned to homogeneous instructional groups within their classes. Within-class ability grouping, such as use of reading or math groups, is the most common form of grouping at the elementary level (McPartland et al., 1987). Complex plans, such as those that involve grouping across grade lines, flexible grouping for particular topics, and part-time grouping, are also occasionally seen in secondary schools. In general, a wider range of grouping plans are used in middle/junior high schools than in senior high schools.

Arguments for and against ability grouping have been essentially similar for 70 years. For example, Turney (1931), summarizing writings of the 1920s, listed advantages and disadvantages of ability grouping. The advantages were as follows:

1. It permits pupils to make progress commensurate with their abilities.
2. It makes possible an adaption of the technique of instruction to the needs of the group.
3. It reduces failures.
4. It helps to maintain interest and incentive, because bright students are not

bored by the participation of the dull.

5. Slower pupils participate more when not eclipsed by those much brighter.

6. *It makes teaching easier.*

7. It makes possible individual instruction to small slow groups.

The following were the disadvantages:

1. Slow pupils need the presence of the able students to stimulate them and encourage them.

2. A stigma is attached to low sections, operating to discourage the pupils in these sections.

3. Teachers are unable, or do not have time, to differentiate the work for different levels of ability.

4. Teachers object to the slower groups.

A research symposium, school board meeting, or PTA meeting on the topic of ability grouping in 1990 is likely to bring up much the same arguments on both sides, with two important additions: the argument that ability grouping discriminates against minority and lower-class students (e.g., Braddock, 1990; Rosenbaum, 1976), and the argument that students in the low tracks receive a lower pace and lower quality of instruction than do students in the higher tracks (e.g., Gamoran, 1989; Oakes, 1985).

In essence, the argument in favor of ability grouping is that it will allow teachers to adapt instruction to the needs of a diverse student body and give them an opportunity to provide more difficult material to high achievers and more support to low achievers. The challenge and stimulation of other high achievers are believed to be beneficial to high achievers (see Feldhusen, 1989). Arguments opposed to ability grouping focus primarily on the perceived damage to low achievers, who receive a slower pace and lower quality of instruction, have teachers who are less experienced or able and who do not want to teach low-track classes, face low expectations for performance, and have few positive behavioral models (e.g., Gamoran, 1989; Oakes, 1985; Persell, 1977; Rosenbaum, 1980). Because of the demoralization, low expectations, and poor behavioral models, students in the low tracks are believed to be more prone to delinquency, absenteeism, dropout, and other social problems (Crespo & Michelna, 1981; Wiatrowski, Hansell, Massey, & Wilson, 1982). With few college-bound peers, students in low tracks have been found to be less likely to attend college than other students (Gamoran, 1987). Ability grouping is perceived to perpetuate social class and racial inequities because lower class and minority students are disproportionally represented in the lower tracks. Ability grouping is often considered to be a major factor in the development of elite and under-class groups in society (Persell, 1977; Rosenbaum, 1980). Perhaps most important, tracking is believed to work against egalitarian, democratic ideals by sorting students into categories from which escape is difficult or impossible.

There are important differences between the pro-grouping and anti-grouping positions that go beyond the arguments themselves. Arguments in favor of ability grouping focus on *effectiveness*, saying in effect that as distasteful as grouping may be, it so enhances the learning of students (particularly but not only high achievers) that its use is necessary. In contrast, arguments opposed to grouping focus at least as much on *equity* as on effectiveness and on democratic values as much as on outcomes. In one sense, then, the burden of proof is on those who favor grouping, for if grouping is not found to be clearly more effective than heterogeneous placement, none of the pro-grouping arguments apply. The same is not true of anti-grouping arguments, which provide a rationale for abolishing grouping that would be plausible even if grouping were found to have no adverse effect on achievement.

Research on the achievement effects of ability grouping has taken two broad forms. One type of research compares the achievement gains of students who are in one or another form of grouping to those of students in ungrouped, heterogeneous placements. Another type of research compares the achievement gains made by students in high-ability groups to those made by students in the low groups.

Reviews of the grouping versus nongrouping literature have consistently shown that grouping has little or no impact on overall student achievement in elementary and secondary schools (e.g., Borg, 1965; Esposito, 1973; Findley & Bryan, 1971; Good & Marshall, 1984; Heathers, 1969; Kulik & Kulik, 1982). Primarily on the basis of his own empirical research, Borg (1965) claimed that ability grouping had a slight positive effect on the achievement of high achievers and a slight negative effect on low achievers, but Kulik and Kulik (1987) found no such trend.

In contrast, researchers who have compared gains made by students in different tracks have generally concluded that controlling for ability level, socioeconomic status, and other control variables, being in the top track accelerates achievement and being in the low track significantly reduces achievement (Alexander, Cook, & McDill, 1978; Dar & Resh, 1986; Gamoran & Berends, 1987; Gamoran & Mare, 1989; Oakes, 1982; Persell, 1977; Sorensen & Hallinan, 1986). In fact, many researchers and theorists in the sociological tradition maintain that tracking is a principal source of social inequality in society and that it causes or greatly magnifies differences along lines of class and ethnicity (e.g., Braddock, 1990; Jones, Erickson, & Crowell, 1972; Schafer & Olexa, 1971; Vanfossen, Jones, & Spade, 1987).

One area of research has investigated the quality of instruction offered to students in high- and low-ability groups, usually concluding that low-ability group classes receive instruction that is significantly lower in quality than that received by students in high-track classes (e.g., Evertson, 1982; Gamoran, 1989; Oakes, 1985; Trimble & Sinclair, 1987). However, it is difficult to compare "quality of instruction" in high- and low-track classes. For example, teachers typically cover less material in low-track classes (e.g., Oakes, 1985). Is this an indication of poor quality of instruction or an appropriate pace of instruction? Students in low-track classes are more off-task than those in high-track classes (e.g., Evertson, 1982). Is this due to the poor behavioral models and low expectations in the low-track classes, or would low achievers be more off-task than high achievers in any grouping arrangement? Evidence that low-track classes are often taught by less experienced or less qualified teachers or that they manifest other objective indicators of lower-quality instruction could justify the conclusion that regardless of measurable effects on learning, students in the lower tracks do not receive equal treatment, but such evidence is rare.

In addition to synthesizing research on overall effects of ability grouping on the achievement of high-average- and low-achieving secondary students, this review will attempt to reconcile research comparing achievement gains in different tracks with research comparing grouped and ungrouped settings.

Review Methods

This review uses a procedure called "best-evidence synthesis" (Slavin, 1986), which incorporates the best features of meta-analytic and traditional reviews. Best-evidence syntheses specify clear, well-justified methodological and substantive criteria for inclusion of studies in the main review and describe individual studies and critical research issues in the depth typical of good-quality narrative reviews. However, whenever possible, effect sizes are used to characterize study outcomes, as in meta-analyses (Glass, McGaw, & Smith, 1981). Systematic literature search procedures, also characteristic of meta-analysis, are similarly applied in best-evidence syntheses.

Criteria for Study Inclusion

The studies on which this review is based had to meet a set of a priori criteria with respect to relevance to the topic and methodological adequacy. First, all studies had to involve comprehensive ability grouping plans that incorporated most or all students in the school. This excludes studies of special programs for the gifted or other high achievers as well as studies of special education, remedial programs, or other special programs for low achievers. Studies of within-class ability grouping are included, but studies of such grouping-related programs as individualized instruction, mastery learning, cooperative learning, and continuous-progress groupings are excluded.

Studies had to be available in English, but otherwise no restrictions were placed on study location or year of publication. Every attempt was made to locate dissertations and other unpublished documents in addition to the published literature.

Methodological requirements for inclusion. Criteria for inclusion of studies in the main review were essentially identical to those used in an earlier review of elementary ability grouping (Slavin, 1987). These were as follows:

1. Ability-grouped classes were compared to heterogeneously grouped classes. This requirement excluded a few studies that correlated "degree of heterogeneity" with achievement gain (e.g., Millman & Johnson, 1964; Wilcox, 1963). Studies that compared achievement gains for students in different tracks but not to heterogeneous classes (e.g., Alexander et al., 1978) were excluded from the main review but are discussed in a separate section.

2. Achievement data from standardized or teacher-made tests were presented. This excluded many anecdotal reports and studies that used grades as the dependent measure. Teacher-made tests, used in a very small number of studies, were accepted only if there was evidence that they were designed to assess objectives taught in all classes.

3. Initial comparability of samples was established by use of random assignment or matching of students or classes. When individual students in intact schools or classes were matched, evidence had to be presented that the intact groups were comparable.

4. Ability grouping had to be in place for at least a semester.

5. At least three ability-grouped and three control classes were involved.

The criteria outlined above excluded very few studies comparing comprehensive ability grouping plans to heterogeneous placements. Every study located that satisfied criteria 1, 2, and 3 also satisfied criteria 4 and 5. Excluding studies of special programs for high achievers (e.g., Atkinson & O'Connor, 1963), all but two of the studies included in meta-analyses by Kulik and Kulik (1982, 1987) were also included in the present review. The exceptions were a study by Adamson (1971) that had substantial IQ differences favoring the ability-grouped school and one by Wilcox (1963) that compared more and less heterogeneous tracked classes.

One major category of studies included in the present review but excluded by the Kuliks includes studies that did not present data from which effect scores could be computed (e.g., Borg, 1965; Ferri, 1971; Lovell, 1960; Postlethwaite & Denton, 1978). These studies are discussed in terms of the direction and statistical significance of their findings.

Literature Search Procedures

The studies included here were located in an extensive search. Principal sources included the Education Resources Information Center (ERIC), *Dissertation Abstracts International*, and citations made in other reviews, meta-analyses, and primary sources. Every attempt was made to obtain a complete set of published and unpublished studies that met the criteria outlined above.

Computation of Effect Sizes

Effect sizes were generally computed as the difference between the experimental and control means divided by the control group's standard deviation (Glass et al., 1981). In ability grouping literature, the heterogeneous group is almost always considered the control group, and this convention is followed in the present article; positive effect sizes are ones that favored ability grouping, whereas negative effect sizes indicated higher means in the heterogeneous groups. The standard deviation of the heterogeneous group is also preferred as the denominator because of the possibility that ability grouping may alter the distribution of scores. However, when means or standard deviations were omitted in studies that otherwise met the inclusion criteria, effect sizes were estimated when possible from *t*s, *F*s, exact *p* values, sums of squares in factorial designs, or other information, following procedures described by Glass et al. (1981).

Several of the studies included in this review presented data comparing gain scores without reporting actual pre- or posttest means. Standard deviations of gain scores are typically lower than those of raw scores (to the degree that pre-post correlations exceed +0.5), so effect sizes computed on gain scores are often inflated. If pre-post correlations are known, effect sizes from all scores can be transformed to the scale of posttest values. However, because none of the studies using gain scores also provided pre-post correlations, a pre-post correlation of +0.8 was assumed (following Slavin, 1987). Using a formula from Glass et al. (1981), this correlation produces a multiplier of 0.632, which was used to deflate effect size estimates from gain score data. The purpose of this procedure and others was to attempt to put all effect size estimates in the same metric, the unadjusted standard deviation of the heterogeneous classes. However, because this multiplier is only a rough approximation, effect sizes from studies using gain scores should be interpreted with even more caution than that which is warranted for effect sizes in general.

Another deviation from usual meta-analytic procedure used in the present review involved adjustments of posttest scores for any pretest differences. This was done either by subtracting pretest means from posttests (if the same tests were used), by converting pre- and posttest means to *z* scores and then subtracting (if different tests were used), or by using covariance-adjusted scores. However, even when such adjustments were made, affecting the numerator of the effect size formula, the denominator remained the unadjusted posttest standard deviation.

One effect size is reported for each study (see Bangert-Drowns, 1986). When multiple subsamples, subjects, or tests were used, medians were computed across the data points. For example, if four measures were used with three subgroups (e.g., high, middle, and low achievers), the effect size for the study as a whole would be the median of the 12 (4 X 3) resulting effect sizes. Whenever possible, findings were also broken down by achievement level (high, average, low), and separate effect sizes were computed for each major subject.

In pooling findings across studies, medians rather than means were used, principally to avoid giving too much weight to outliers. However, any measure of central tendency in a meta-analysis or best-evidence synthesis should be interpreted in light of the quality and consistency of the studies from which it was derived, not as a finding in its own right.

Research on Ability Grouping in Secondary Schools

A total of 29 studies of tracking or streaming in secondary schools met the inclusion criteria listed earlier. The studies, their major characteristics, and their findings are listed in Table 1.

The studies listed in Table 1 are organized in three categories according to their research designs. Six studies used random assignment of students to ability-grouped or heterogeneous classes. Nine studies took groups of students; matched them individually on IQ, composite achievement, and other measures; and then assigned one of each matched pair of students to an ability-grouped class and one to a heterogeneous class. The quality of these randomized or matched experimental designs is very high, and the findings of the 15 studies using such designs must be given special weight. The remaining 14 studies investigated existing schools or classrooms that used or did not use ability grouping, and then either selected matched groups of students from within each type of school or used analyses of covariance or other statistical procedures to equate the groups. The difficulty inherent in such designs is that any differences between schools that are systematically related to ability grouping would be confounded with the practice of ability grouping per se. For example, a secondary school that used heterogeneous grouping might have a staff, principal, or community more concerned about equity, affective development, or other goals than would a "matched" school that used ability grouping. However, several of the correlation studies used very large samples and longitudinal designs, and these provide important additional information not obtainable from the typically smaller and shorter experimental studies.

Within each category, studies are listed in descending order of sample size. All other things being equal, therefore, studies near the top of Table 1 should be considered as better evidence of the effects of ability grouping than studies near the end of the table. However, the nature and quality of the studies are discussed in more detail in the following sections.

TABLE 1

					Effect Sizes			
Article	Grades	Location	Sample size	Duration	Design	By achievement	By subject	Total
		Randomized experimental studies						
Marascuilo & McSweeney, 1972	8-9	Berkeley, CA	603 students	2 yrs.	Students randomly assigned to 3-group AG or hetero social studies classes. Compared students on teacher-made and standardized tests.	HI +.14 AV -.37 LO -.43	Social studies -.22	-.22
Drews, 1963	9	Lansing, MI	4 schools, 432 students	1 yr.	Students randomly assigned to 3-group AG or hetero English classes compared on standardized tests.	HI -.16 AV +.01 LO -.01	Reading -.11 Language .00	-.05
Fick, 1963	7	Olathe, KS	1 school, 168 students	1 yr.	Students randomly assigned to 3-group AG or hetero "core" classes. Both classes taught by same teacher. Iowa Tests of Basic Skills used as posttests.	HI +.01 AV .00 LO -.04	Reading -.01 Language .00	-.01
Peterson, 1966	7-8	Chisholm, MN	1 school, 152 students	1 yr.	Students randomly assigned to AG or hetero classes. AG based on composite achievement, grades. Compared on standardized tests.	HI +.05 AV -.44 LO -.06	Reading +0.2 Language -.01 Math -.25 Social Studies -.07	-.04
Ford, 1974	9	New York, NY (Low SES minority)	80 students	1 sem.	Students randomly assigned to 2-group AG or hetero math classes. Same teachers taught both types of classes. Students compared on Metropolitan Achievement Test.		Math (0)	(0)
Bicak, 1962	8	Minneapolis, MN (lab school)	1 school, 75 students	1 sem.	Students randomly assigned to 2-group AG or hetero science classes at university lab school.	HI -.39 LO -.10	Science -.25	-.25
		Matched experimental studies						
Lovell, 1960	10	Panama City, FL	500 students	1 yr.	Matched students assigned to 5-group AG or hetero English, biology, and algebra classes.	HI (+) AV (0) LO (0)	Language (+) Math (0) Biology (0)	(0)

					Effect Sizes			
Article	Grades	Location	Sample size	Duration	Design	By achievement	By subject	Total
Billett, 1928	9	Painesville, OH	408 students	1 yr.	In 3 successive years, matched students assigned to 3-group AG or hetero English classes. Compared gains on standardized tests.	HI -.11 AV +.03 LO +.18	English +.04	+.04
Platz, 1965	9	—[a]	298 students	1 sem.	Matched students assigned to 3-group AG or hetero science classes. Students compared on standardized science test.	HI +.18 AV -.10 LO +.22	Science +.22	+.22
Bailey, 1968	9	St. Louis, MO	255 students	1 yr.	Matched students assigned to 2-group AG or hetero algebra classes. Same teachers taught both types of classes. Students compared on gains on standardized algebra measure.	HI +.18 LO -.24	Math -.03	-.03
Thompson, 1974	11	Suburban, VA	240 students	1 yr.	Compared students in 2 schools, one AG in social studies, one hetero. Students matched. Compared gain scores on teacher-made tests.	HI -.50 HI AV -.47 LO AV -.43 LO -.54	Social studies -.48	-.48
Barton, 1964	9	Rural UT	204 students	1 yr.	Matched students assigned to 4-group AG in math or to hetero English classes, compared on California Achievement Test gains.	HI +22 HI AV -.03 LO AV -.13 LO -.20	Reading +.06 Language -.13	-.04
Willcutt, 1969	7	Bloomington, IN (lab school)	156 students	1 yr.	Matched students assigned to 4-group flexible AG in math or to hetero. Grouping changed 8 times in the year.		Math -.15	-.15
Holy & Sutton, 1930	9	Marion, OH	148 students	1 sem.	Matched students assigned to AG, hetero algebra classes. Same teacher taught all classes.		Math +.28	+.28

					Effect Sizes			
Article	Grades	Location	Sample size	Duration	Design	By achievement	By subject	Total
Martin, 1927	7	New Haven, CT	83 students	1 yr.	Matched students assigned to 3-group AG or hetero.	HI +.12 AV -.06 LO +.23	Reading +.17 Math +.13 Language +.03 Social studies .00 Science -.04	+.10
		Correlational Studies						
Kerchkoff, 1986	5-10	Britain	8,500 students	5 yrs.	Longitudinal study of students throughout Britain who attended streamed or unstreamed secondary schools.		Reading +.02 Math +.03	+.03
Fogelman, Esen, & Tibbenham, 1978	6-10	Britain	5,923 students	4 yrs.	Retrospective study compared students who had been streamed, partially streamed, or heterogeneous schools throughout secondary school, controlling for Grade 5 general ability			
Borg, 1965	6-9 7-10 8-11 9-12	Utah	2,934 students	4 yrs.	Longitudinal study of students in districts using AG compared to students in neighboring district using heterogeneous grouping, controlling for pretests.	HI (0) AV (0) LO (0)	Math (0) Science (0)	(0)
Ferri, 1971	5-6	Britain	28 schools, 1,716 students	2 yrs.	Streamed and nonstreamed schools matched on 7+ (Grade 2) reading, followed 4 years in junior school, 2 years in secondary.	HI (0) AV (0) LO (0)	Math (0) English (0)	(0)
Breidenstine, 1936	7-9	Soudersburg, PA	11 schools, 860 students	1 yr.	Compared to students in 4 AG, 7 hetero schools matched on IQ.		Composite achievement	-.19
Purdom, 1929	9	--a	700 students	1 sem.	Matched students in AG, hetero English and algebra classes compared in achievement	HI -.02 AV -.08 LO +.07	English -.02 Algebra .00	+.01

						Effect Sizes		
Article	**Grades**	**Location**	**Sample Size**	**Duration**	**Design**	**By Achievement**	**By Subject**	**Total**
Postlethwaite & Denton, 1978; Newbold, 1977	5-7	Britain	1 school, 450 students	2 yrs.	Students within one secondary school assigned to streamed or unstreamed halls. Achievement assessed on national examinations.	HI (0) AV(0) LO (0)	Math (0) English (0) Social studies (0) French (0)	(0)
Bachman	7	Portland, OR	15 schools, 23 teachers, 404 students	1 yr.	Math classes in schools using AG compared to hetero classes, controlling for IQ.		Math(0)	(0)
Kline, 1964	9-12	St. Louis, MO	4 schools	4 yrs.	Retrospective study of successive cohorts of students, one in 3- or 4-group AF, one hetero, in 4 schools. Compared on standardized tests after 4 years of AG or hetero placement.	V.HI -.02 HI +.08 AV .00 LO -.02	Reading -.05 Language +.07 Math +.01	+.01
Stoakes, 1964	7	Cedar Rapids, IA	3 schools	1 yr.	Matched mentally advanced and slow-learning students compared in schools using AG or hetero assignment. Compared on standardized tests.	HI (0) LO (0)	Reading (0) English (0) Math (0)	(0)
Martin, 1959	6-8	Nashville, TN	3 schools	2 yrs.	Retrospective study compared gains on Standford Achievement Tests for 2 AG and 1 hetero school from Grades 6-8.	HI (0) AV (0) LO (0)	Reading (0) Language (0) Math (0)	(0)
Chiotti, 1961	9	Issaquah, WA	3 schools	1 yr.	Matched students in 3-group AG and hetero schools compared in math achievement.	HI +.14 AV +.06 LO +.35	Math +.18	+.18
Fowlkes, 1931	7	Glendale, CA	2 schools	1 sem.	Students in school using 3-group AG based on IQ matched with students in hetero school. Compared gains on Stanford Achievement Tests.	HI -.45 AV -.18 LO -.05	Reading -.04 Language -.17 Math -.17 Social Studies -.21	-.20
Cochran, 1961	8	Kalamazoo, MI	1 school	1 yr.	Compared students grouped separately for English, math, to previous year (hetero) students matched in IQ, age, sex, achievement.	HI (0) AV (0) LO (0)	Math (0) English (0)	(0)

Overall Findings

Across the 29 studies listed in Table 1, the effects of ability grouping on student achievement are essentially zero. The median effect size (ES) for the 20 studies from which effect sizes could be estimated was -.02, and none of the 9 additional studies found statistically significant effects. Counting the studies with nonsignificant differences as though they had effect sizes of .00, the median effect size for all 29 studies would be .00. Results from the 15 randomized and matched experimental studies were not much different; the median effect size was -.06 for the 13 studies from which effect sizes could be estimated. In 9 of these 13 studies (including all 5 of the randomized studies) results favored the heterogeneous groups, but these effects are mostly very small.

There are a few consistent patterns in the study findings. Most of the studies involved Grades 7-9, with ninth graders sometimes in junior high schools and sometimes in senior high schools. No apparent trend is discernible within this range. Above the ninth grade the evidence is too sparse for firm conclusions. Lovell (1960) found that high achieving tenth graders performed significantly better in ability-grouped English classes, but there were no effects in biology or algebra and no effects for average or low achievers. In a 4-year study of students in Grades 9-12, Borg (1965) found significant positive effects of ability grouping for average and low achievers in math but no differences in science or for high achievers. Cohorts followed from Grades 7-10 and 8-11 showed no significant differences on any measure for any ability level. In contrast, Thompson (1974), in a study of 11th-grade social studies, found the largest effects favoring heterogeneous grouping (ES = -.48), whereas Kline (1964), in another 4-year study of students in Grades 9-12, found no differences.

Twelve of the 29 studies tracked students for all subjects according to one composite ability or achievement measure. The remaining 17 studies grouped students on the basis of performance in one or more specific subjects. However, there were no differences in the outcomes of the different forms of ability grouping. In addition, there were no consistent patterns in terms of the number of ability groups to which students were assigned (the great majority of studies used 3). Study duration had no apparent impact on outcome. Studies that used adjusted gain scores produced the same effects as other studies, and the use of the adjustment of gain scores described above made no difference in outcomes.

There was no discernible pattern of findings with respect to different subjects, with one possible exception. Studies by Marascuilo and McSweeney (1972), Thompson (1974), and Fowlkes (1931) found relatively strong effects favoring heterogeneous grouping in social studies, and three additional studies by Peterson (1966), Martin (1927), and Postlethwaite and Denton (1978) found no differences or slight effects in the same direction. This is not enough evidence to conclusively point to a positive effect of heterogeneous grouping in social studies, but it is important to note that all three of the randomized or matched experimental studies found differences in this direction.

There were no consistent effects according to study location. All four of the British studies found no differences between streamed and unstreamed classes; a large, longitudinal Swedish study by Svensson (1962), not shown in Table 1 because it lacked adequate evidence of initial equality, also found no differences between streamed and unstreamed classes. Urban, suburban, and rural schools had similar outcomes. The one study that involved large numbers of minority students, a randomized experiment in a

New York City high school by Ford (1974), found no differences between ability-grouped and heterogeneous math classes.

Studies conducted before 1950 were no more likely than more recent studies to find achievement differences. On this topic, it is interesting to note that experimental-control studies of ability grouping have not been done in recent years. The only study of the 1980s, by Kerckhoff (1986), was done by a sociologist who focused his attention on differences between students in different streams. This study is described in more detail below. Otherwise, the most recent experimental-control comparisons were done in the early 1970s.

Differential Effects According to Achievement Levels

One of the most important questions about ability grouping in secondary schools concerns the degree to which it differentially affects students at different achievement levels. As noted earlier, many researchers and reviewers, particularly those working in the sociological tradition, have emphasized the *relative* impact of grouping for different groups of students far more than the average effect for all students.

Twenty-one of the 29 studies presented in Table 1 presented data on the effects of ability grouping on students of different ability levels. Most studies divided their samples into three categories (high, average, and low achievers), but some used two or four categories.

Across the 15 studies from which effect sizes could be computed, the median effect size was +.01 for high achievers, -.08 for average achievers, and -.02 for low achievers. Effects of this size are indistinguishable from zero, and if all the nonsignificant differences found in studies from which effect sizes could not be computed are counted as effect sizes of .00, the median effect size for each level of student becomes .00. In addition, only one of seven studies from which effect sizes could not be computed (Lovell, 1960) found significantly positive effects of ability grouping for high achievers, and none of these studies found significant effects in either direction for average and low achievers. The randomized and matched experimental studies provided slightly more support for the idea that ability grouping has a differential effect; the median effect sizes for high, average, and low achievers were +.05, -.10, and -.06, respectively. It is interesting to note that the study by Borg (1965), which is often cited to support the differential effect of ability grouping on students of different ability levels, in fact provides very weak support for this phenomenon. Across two measures given to members of four 4-year cohorts that principally included secondary years, significant effects favoring ability grouping were found for high achievers in one of eight comparisons, for average achievers in three of eight, and for low achievers in one of eight. Only in a cohort that included Grades 4 to 7 were there significant effects favoring heterogeneous grouping for low achievers.

It might be expected that differential effects of track placement would build over time and that longitudinal studies would show more of a differential impact than 1-year studies. The one multiyear randomized study, by Marascuilo and McSweeney (1972), did find that over a 2-year period, students in the top social studies classes gained slightly more than similar students in heterogeneous classes (ES = +.14), whereas middle (ES = -.37), and low (ES = -.43) groups gained significantly less than their ungrouped counterparts. However, across seven multiyear correlational studies of up to 5 years' duration, not one found a clear pattern of differential effects.

A few studies provided additional information on differential effects of ability grouping by investigating effects of grouping on high or low achievers only. For example, Torgelson (1963) randomly assigned low achieving students in Grades 7-9 to homogeneous or heterogeneous classes. Across several performance measures, the median effect size was +.13 (nonsignificantly favoring ability grouping). Similarly, Borg and Prpich (1966) randomly assigned low achieving 10th graders to ability-grouped or heterogeneous English classes and found that there were no differences in one cohort. In a second cohort, differences favoring ability grouping on a writing measure were found, but there were no differences on eight other measures.

Studies of ability grouping of high achievers are difficult to distinguish from studies of special programs for the gifted. Well-designed studies of programs for the gifted generally find few effects of separate programs for high achievers unless the programs include acceleration (exposure to material usually taught at a higher grade level) (Fox, 1979; Kulik & Kulik, 1984). That is, grouping per se has little effect on the achievement of high achievers. An outstanding illustration of this is a dissertation by Mikkelson (1962), who randomly assigned high achieving seventh and eighth graders to ability-grouped or heterogeneous math classes. The seventh grade homogeneous classes were given enrichment, but the eighth graders were accelerated, skipping to ninth grade algebra. No effects were found for the seventh graders. The accelerated eighth graders, of course, did substantially better than similar students who were not accelerated on an algebra test, and they did no worse on a test of eighth grade math.

Taken together, research comparing ability-grouped to heterogeneous placements provides little support for the proposition that high achievers gain from grouping whereas low achievers lose. However, there is an important limitation to this conclusion. In most of the studies that compared tracked to untracked grouping plans (including all of the randomized and matched experimental studies), tracked students took different levels of the same courses (e.g., high, average, or low sections of Algebra I). Yet much of the practical impact of tracking, particularly at the senior high school level, is on determining the nature and number of courses taken in a given area. The experimental studies do not compare students in Algebra I to those in Math 9, or students who take 4 years of math to those who take 2. The conclusions drawn in this section are limited, therefore, to the effects of between-class grouping *within the same courses*, and should not be read as indicating a lack of differential effects of tracking as it affects course selection and course requirements.

Other Forms of Ability Grouping

The studies discussed above and summarized in Table 1 evaluated the most common forms of ability grouping in secondary schools—full-time, between-class ability grouping for one or more subjects. However, a few studies have evaluated other grouping plans.

The most widely used form of grouping in elementary schools, within-class ability grouping, has also been evaluated in a few studies involving middle and junior high schools. Campbell (1965) compared the use of three math groups within the class to heterogeneous assignment in two Kansas City junior high schools. There were no differences between the two programs in achievement. Harrah (1956) compared five types of within-class grouping in Grades 7-9 in West Virginia and found ability grouping to be no more successful than other grouping methods. Note that these findings conflict with those of studies of within-class ability grouping in mathematics in the upper elementary grades, which tended to support the use of math groups (Slavin, 1987).

Vakos (1969) evaluated the use of a combination of heterogeneous and homogeneous instruction in 11th grade social studies classes in Minneapolis. Students were grouped by ability 2 days each week, but heterogeneously grouped the other 3 days. No achievement differences were found. Zweibelson, Bahnmuller, and Lyman (1965) evaluated a similar mixed approach to teaching ninth grade social studies in New Rochelle, New York, and also found no achievement differences. Chiotti (1961) compared a flexible plan for grouping junior high school students across grade lines for mathematics to both ability-grouped and heterogeneous grouping plans, and again found no differences in achievement. A cross-grade grouping arrangement similar to the Joplin Plan (Slavin, 1987) was compared to within-class grouping in reading by Chismar (1971) in Grades 4-8. Significantly positive effects of this program were found in Grades 4 and 7 but not 5, 6, and 8.

Reconciling Track/No-Track and High-Track/Low-Track Studies

As noted earlier, two very different traditions of research have dominated research on ability grouping. One involves comparisons of ability-grouped to heterogeneous placements. The other involves comparisons of the progress made by students in different ability groups or tracks. Whereas there has been little experimental research comparing ability-grouped to heterogeneous placements since the early 1970s, research comparing the achievement of students in different tracks largely began in the 1970s and continues to the present.

The findings of high-track/low-track studies of ability grouping conflict with those emphasized in this review in that they generally find that even after controlling for IQ, socioeconomic status, pretests, and other measures, students in high tracks gain significantly more in achievement than do students in low tracks, especially in mathematics (see Gamoran & Berends, 1987, for a review). How can these findings be reconciled with those of the experimental studies?

One important difference between experimental and correlational studies of ability grouping is that, as mentioned earlier, correlational studies (especially at the senior high school level) often include not only the effects of being in a high, average, or low class, but also the effects of differential course taking. Students in academic tracks may score better than those in general or vocational classes because they take more courses or more advanced courses. The experimental studies comparing grouped and ungrouped classes are all studies of grouping per se, holding course taking and other factors constant. The correlational studies examine tracking as it is in practice, where track placement implies differences in course requirements, course taking patterns, and so on. Also, experimental track versus no-track studies are rare beyond the ninth grade, whereas most correlational studies comparing students in high versus low tracks involve senior high schools. The lack of track versus no-track studies at the senior high school level is hardly surprising given the nearly universal use of some form of tracking at that level. However, tracking usually has a different meaning in senior than in junior high school. Whereas junior high school tracking mostly involves different levels of courses (e.g., high English vs. low English), senior high tracking is more likely to involve completely different patterns of coursework (e.g., metal shop vs.

French III). Also, the problem of dropouts becomes serious in senior high school; a study of 12th graders unavoidably excludes the students who may have suffered most from being in the low track and left school (see Gamoran, 1987). This could reduce observed differences between high- and low-track students.

There is limited evidence, however, that differences in course taking or grade level account for the different conclusions of the track/no-track and high-track/low-track studies. Four-year longitudinal studies in U.S. senior high schools by Kline (1964) and Borg (1965) found no differential effects of track placement for high, average, and low achievers (as compared to similar students in untracked placements). Presumably, course-taking patterns in these senior high school studies varied by track. A correlational study by Alexander and Cook (1982) found that although taking more courses in senior high school did increase achievement (controlling for background factors), different course-taking patterns in different tracks did not account for track differences in achievement. Gamoran (1987) found that track effects on math and science achievement were explained in part by the fact that students in the academic tracks take more math and science courses and, in particular, more *advanced* courses in these areas. However, no such patterns were seen on reading, vocabulary, writing, or civics achievement measures. Gamoran noted the difficulty of disentangling track and course taking, which are highly correlated in math and science (and, of course, both track and course taking are strongly correlated with ability, socioeconomic status, and other factors). It is certainly logical to expect correlational studies of senior high school tracking to find different effects of different track placements because of different course-taking patterns, but because of confounding of tracking, course-taking, and student background factors, that is difficult to determine conclusively.

Another likely explanation for different findings of track/no-track and high-track/low-track studies involves the difficulty of statistically controlling for large differences. Students in higher tracks tend to achieve at much higher levels than those in lower tracks (both before and after taking secondary courses), and statistically controlling for these differences is probably not sufficient to completely remove the influence of ability or prior performance on later achievement. Further, studies in higher tracks are also likely to be higher in such attributes as motivation, internal locus of control, academic self-esteem, and effort, factors that are not likely to be controlled in correlational studies.

To understand the difficulty of controlling for large initial differences between students, imagine an experiment in which a new instructional method was to be evaluated. The experimenter selects a group of students who have high test scores and high IQ scores and are nominated by their teachers as being hard working, motivated, and college material. This becomes the experimental group, and the remaining students serve as the control group. To control for the differences between the groups, prior composite achievement and socioeconomic status are used as covariates or control variables.

In such an experiment, no one would doubt that regardless of the true effectiveness of the innovative treatment, the experimental group would score far better than the control group, even controlling for prior achievement and socioeconomic status. No journal or dissertation committee would accept such a study. Yet this "experiment" is essentially what is being done when researchers compare students in different tracks. When there are significant pretest differences, use of statistical controls through analysis of covariance or regression are considered inadequate to equate the groups. Most often, the statistical controls will undercontrol for true differences (Lord, 1960; Reichardt, 1979). Yet high- and low-track students usually differ in pretests or IQ by one to two standard deviations, an enormous systematic difference for which no statistical procedure can adequately control.

The only study that compared both tracked to untracked schools and high-track to low-track students was a 5-year longitudinal study by Kerckhoff (1986) in Britain. This study illustrates the problem of controlling for large differences. For example, in mathematics, boys in the high track of three-group ability grouping programs gained about 11 *z* score points from a test given at age 11 to one given at age 16, whereas students in a remedial track gained 18 *z* score points. Yet the regression coefficient comparing the high-track to ungrouped students was +2.34, indicating performance about 42% of a standard deviation above "predicted" performance. In contrast, the remedial-track boys had a regression coefficient (in comparison to ungrouped students) of -.72, indicating performance about 13% of a standard deviation below "predicted" performance, despite the fact that the remedial students actually gained more than the top-track students. The reason for this is that the remedial students started out (at age 11) scoring 1.64 standard deviations below the ungrouped students, whereas top-track students started out 1.02 standard deviations above the ungrouped students, a total difference between top-track and remedial students of 2.66. No regression or analysis of covariance can adequately control for such large pretest differences. Because of unreliability in the measures and less-than-perfect within-group correlations of pre- and posttests, "predicted" scores based on pretests and other covariates will (other things being equal) be too low for high achievers and too high for low achievers.

Another factor that can contribute to overestimates of the effects of curriculum track on achievement in studies lacking heterogeneous comparison groups is fan spread. Put simply, high achievers usually gain more per year than do low achievers, so over time the gap between high and low achievers grows. This increasing gap cannot be unambiguously ascribed to ability grouping or other school practices, as it occurs under virtually all circumstances. A student who is performing at the 16th percentile in the sixth grade and is still at the 16th percentile in 12th grade will be further "behind" the 12th grade mean in grade equivalents, for example (Coleman & Karweit, 1972).

An additional factor that can contribute to spurious findings indicating a benefit of being in the high track is that factors other than test scores factor into placement decisions. For example, a study by Balow (1964) found that on math tests not used for group placement, there was enormous overlap between students in supposedly homogeneous seventh-grade math classes. More than 72% of the students scored between the lowest score in the top group and the highest score in the bottom group. Among these students in the "area of overlap," students who were in the top group gained the most in math achievement over the course of the year, whereas those in the low group gained the least.

On its surface this study provides support to the "self-fulfilling prophecy" argument. Yet consider what is going on. Imagine two students with identical scores, one assigned to the high group and one to the low group. Why were they so assigned? Random error is a possibility, but all the systematic possibilities weigh in the direction of higher performance for the student assigned to the high group. Because teacher judgment was involved, teachers may have accurate knowledge of student motivation, self-esteem, behavior, or other factors to enable them to predict who will do well and who will not. The actual assignments were done on different tests than those used in the Balow study; it is likely that students who scored low on Balow's pretests but were put in the high groups scored high on the test used for placement, and then regressed to a higher mean on Balow's posttest.

What this discussion is meant to convey is not that different tracks do or do not have a differential impact on student achievement, but that comparisons of students in existing tracks cannot tell us one way or another. To learn about the differential impacts of track placement, there are two types of research that might be done. One would be to randomly assign students at the margin to different tracks, something that has never been done. The other is to compare similar students randomly assigned to ability-grouped or ungrouped systems. This has been done several times, and, as noted earlier in this review, there is no clear trend indicating that students in high-track classes learn any more than high achieving students in heterogeneous classes, or that students in low-track classes learn any less than low achieving students in heterogeneous classes.

Why Is Ability Grouping Ineffective?

The evidence summarized in Table 1 and discussed in this review is generally consistent with the conclusions of earlier reviews comparing homogeneous and heterogeneous grouping (e.g., Kulik & Kulik, 1982, 1987; Noland, 1985), but runs counter to two quite different kinds of "common sense." On one hand, it is surprising to find that assignment to the low-ability group is not detrimental to student learning. A substantial literature has indicated the low quality of instruction in low groups (e.g., Evertson, 1982; Gamoran, 1989; Oakes, 1985), and a related body of research has documented the negative impact of ability grouping on the motivations and self-esteem of students assigned to low groups (e.g., Cottle, 1974; Schafer & Olexa, 1971; Trimble & Sinclair, 1987). How can the effect of ability grouping on low achieving students be zero, as this review concludes?

On the other hand, another kind of "common sense" would argue that, at least in certain subjects, ability grouping is imperative in secondary schools. How can an 8th grade math teacher teach a class composed of students who are fully ready for algebra and students who are still not firm in subtraction and multiplication? How does an English teacher teach literature and writing to a class in which reading levels range from 3rd to 12th grade? Yet study after study, including randomized experiments of a quality rarely seen in educational research, finds no positive effect of ability grouping in any subject or at any grade level, even for the high achievers most widely assumed to benefit from grouping.

The present review cannot provide definitive answers to these questions. However, it is worthwhile to speculate on them.

One possibility is that the standardized tests used in virtually all of the studies discussed in this review are too insensitive to pick up effects of grouping. This seems particularly plausible in looking at tests of reading, because reading has not generally been taught as such in secondary schools. However, standardized tests of mathematics do have a great deal of face validity and curricular relevance, and these show no more consistent a pattern of outcomes. Marascuilo & McSweeney (1972) used both teacher-made and standardized measures of social studies achievement and found similar results with each.

Another possibility is that it simply does not matter whom students sit next to in a secondary class. Secondary teachers use a very narrow range of teaching methods, overwhelmingly using some form of lecture or discussion (Goodlad, 1983). In this setting, the direct impact of students on one another may be minimal. If this is so, then any impacts of ability grouping on students would have to be mediated by teacher characteristics or behaviors or by student perceptions and motivations.

Studies contrasting teaching behaviors in high- and low-track classes usually find that the low tracks have a slower pace of instruction and lower time on-task (e.g., Evertson, 1982; Oakes, 1982). Yet, as noted earlier, the meaning and impact of these differences are not self-evident. It may be that a slower pace of instruction is appropriate with lower-achieving students, or that pace is relatively unimportant because a higher pace with lower mastery is essentially equivalent to a lower pace with higher mastery. Higher time on-task should certainly be related to higher achievement (Brophy & Good, 1986), but the comparisons of time on task between high and low tracks are misleading. What would be important to compare is time on task *for low achievers* in homogeneous and heterogeneous classes, because low achievers may simply be off-task more than high achievers regardless of their class placement. In this regard, it is important to note that Evertson, Sanford, and Emmer (1981) found time on-task to be lower in extremely heterogeneous junior high school classes than in less heterogeneous ones because teachers had difficulty managing the more heterogeneous classes.

The lesson to be drawn from research on ability grouping may be that unless teaching methods are systematically changed, school organization has little impact on student achievement. This conclusion would be consistent with the equally puzzling finding that substantial reductions in class size have little impact on achievement (Slavin, 1989); if teachers continue to use some form of lecture/discussion/seatwork/quiz, then it may matter very little in the aggregate which or how many students the teachers are facing. In contrast, forms of ability grouping that were found to make a difference in the upper elementary grades—the Joplin Plan (cross-grade grouping in reading to allow for whole-class instruction) and within-class grouping in mathematics (Slavin, 1987)—both significantly change time allocations and instructional activities within the classroom.

Alternatives to Ability Grouping

If the effects of ability grouping on student achievement are zero, then there is little reason to maintain the practice. As noted earlier in this article, arguments in favor of ability grouping depend on assumptions about the effectiveness of grouping, at least for high achievers. In the absence of any evidence of effectiveness, these arguments cannot be sustained.

Yet there is also no evidence that simply moving away from traditional ability grouping practices will in itself enhance student achievement, and there are legitimate concerns expressed by teachers and others about the practical difficulties of teaching extremely heterogeneous classes at the secondary level. How can schools moving away from traditional ability grouping use this opportunity to contribute to student achievement?

One alternative to ability grouping often proposed (e.g., Oakes, 1985) is the use of cooperative learning methods, which involve students working in small, heterogeneous learning groups. Research on cooperative learning consistently finds positive effects of these methods if they incorporate two major elements: group goals and individual accountability (Slavin, 1990). That is, the cooperating groups must be rewarded or recognized on the basis of the sum or average of individual learning performances. Cooperative learning methods of this kind have been used successfully at all grade levels, but there is less research on them in Grades 10-12 than in Grades 2-9 (see Newmann & Thompson, 1987). Cooperative learning methods have also had consistently positive impacts on such outcomes as self-esteem, race relations, acceptance of mainstreamed academically handicapped students, and ability to work cooperatively (Slavin, 1990).

One category of cooperative learning methods may be particularly useful in middle schools moving toward heterogeneous class assignments. These methods are Cooperative Integrated Reading and Composition (Stevens, Madden, Slavin, & Farnish, 1987) and Team Assisted Individualization—Mathematics (Slavin & Karweit, 1985; Slavin, Madden, & Leavey, 1984). Both of these methods are designed to accommodate a wide range of student performance levels in one classroom, using both homogeneous and heterogeneous within-class groupings. These programs have been successfully researched in Grades 3-6 but are often used up to the eighth grade level.

Other alternatives to between-class ability grouping have also been found to be successful in the upper elementary grades (see Slavin, 1987) and could probably be effective in middle schools as well. These include within-class ability grouping in mathematics (e.g., teaching two or three math groups within a heterogeneous class) and the Joplin Plan in reading. The Joplin Plan involves regrouping students for reading across grade levels but according to reading level, so that no within-class reading groups are necessary. However, although these alternatives to between-class grouping are promising because of their success in the upper elementary grades, the few studies of within-class ability grouping at the junior high school level have not found this practice to be effective (Campbell, 1965; Harrah, 1956), and the one middle school study of the Joplin Plan found only inconsistent positive effects (Chismar, 1971).(For descriptions of secondary school schools implementing alternatives to traditional ability grouping, see Slavin, Braddock, Hall, & Petza, 1989).

Limitations of This Review

It is important to note several limitations of the present review. Perhaps the most important is that in none of the studies reviewed here were there systematic observations made of teaching and learning. Observational studies and outcome studies have proceeded on parallel tracks; it would be important to be able to relate evidence of outcomes to changes in teacher behaviors or classroom characteristics. In particular, it would be important to know the degree to which teachers in ability-grouped schools actually differentiate instruction. For example, are teachers of high-track classes more likely to provide enrichment (e.g., greater depth on the same objectives) or acceleration (e.g., coverage of more material usually taught at a later grade level)? How do teachers of low-track classes adapt instruction to the needs of their students? How do teachers of untracked, heterogeneous classes accommodate the wide range of performance levels in their classes? What level and pace of instruction is provided in untracked, heterogeneous classes? Most important, how do variations from teacher to teacher in instructional behaviors in high, low, and heterogeneous classes relate to the outcomes of ability grouping for students of different ability levels?

Another limitation, mentioned earlier, is that almost all studies reviewed here used standardized tests of unknown relationship to what was actually taught. It may be, for example, that positive effects of ability grouping for high achievers could be missed by standardized tests because what these students are getting is enrichment or higher-order skills not assessed on the standardized measures, or that negative effects for low achievers are missed because teachers of low-track classes are hammering away at the minimum skills that are assessed on the standardized tests but ignoring other content. Future research on ability grouping needs to closely examine possible outcomes of grouping on more broadly based and sensitive measures.

A third limitation is the age of most of the studies reviewed. It is possible that schools, students, or ability grouping have changed enough since the 1960s and 1970s to make conclusions from these and older studies tenuous.

As noted earlier, the results reported in this review mainly concern the effects of grouping per se, with little regard for the effects of tracking on such factors as course taking. Effects of tracking on differential course taking are most important in senior high schools. There is a need for additional research comparing tracked to untracked situations at the senior high school level, particularly research designed to disentangle the effects of tracking from those of differential course taking.

In addition, it would add greatly to the understanding of ability grouping in secondary schools to have evaluations or even descriptions of a wider range of alternatives to traditional ability groupings. The few studies of within-class grouping, cross-grade grouping, and flexible grouping plans are not nearly adequate to explore alternatives. Cooperative learning, often proposed as an alternative to ability grouping, has frequently been found to increase student achievement in ability-grouped as well as ungrouped secondary classes (Newmann & Thompson, 1987; Slavin, 1990), yet no study has compared cooperative learning in heterogeneous classes to traditional instruction in homogeneous ones. Descriptions of creative alternatives to ability grouping currently exist only at the anecdotal level (Slavin et al., 1989).

Conclusions

Although there are limitations to the scope of this review and to the studies on which it is based, there are several conclusions that can be advanced with some confidence. These are as follows:

> 1. Comprehensive between-class ability grouping plans have little or no effect on the achievement of secondary students, at least as measured by standardized tests. This conclusion is most strongly supported in Grades 7-9, but the more limited evidence that does exist from studies in Grades 10-12 also fails to support any effect of ability grouping.
>
> 2. Different forms of ability grouping are equally ineffective.
>
> 3. Ability grouping is equally ineffective in all subjects, except that there may be a negative effect of ability grouping in social studies.
>
> 4. Assigning students to different levels of the same course has no consistent positive or negative effects on students of high, average, or low ability.

For the narrow but extremely important purpose of determining the impact of ability grouping on standardized achievement measures, the studies reviewed here are exemplary. Six randomly assigned individual students to ability-grouped or heterogeneous classes, and nine more individually matched students and then assigned them to one or the other grouping plan. Many of the studies followed students for 2 or more years. If there had been any true effect of ability grouping on student achievement, this set of studies would surely have detected it.

For practitioners, the findings summarized above mean that decisions about whether or not to ability group must be made on bases other than likely effects on achievement. Given the antidemocratic, antiegalitarian nature of ability grouping, the burden of proof should be on those who would group rather than those who favor heterogeneous grouping, and in the absence of evidence that grouping is beneficial, it is hard to justify continuation of the practice. The possibility that students in the low groups are at risk for delinquency, dropout, and other social problems (e.g., Rosenbaum, 1980) should also weigh against the use of ability grouping. Yet schools and districts moving toward heterogeneous grouping have little basis for expecting that abolishing ability grouping will in itself significantly accelerate student achievement unless they also undertake changes in curriculum or instruction likely to improve actual teaching.

There is much research still to be done to understand the effects of ability grouping in secondary schools on student achievement. Studies using more sensitive achievement measures, studies of grouping at Grades 10-12, studies of a broader range of alternatives to grouping, and studies relating observations to outcomes of grouping are areas of particular need. Enough research has been done comparing the effects of ability grouping on standardized achievement tests for students assigned to high, middle, and low tracks, at least up through the ninth grade. It is time to move beyond these simple comparisons to consider more fully how secondary schools can adapt instruction to the needs of a heterogeneous student body.

Author's Note: This paper was written under funding from the Office of Educational Research and Improvement, U.S. Department of Education, as part of a subcontract with the National Center on Effective Secondary Schools at the University of Wisconsin. However, any opinions expressed are those of the author and do not represent OERI positions or policy. I would like to thank Fred Newmann, Jeannie Oakes, Adam Gamoran, Harry Passow, and Alan Kerckhoff for their comments on an earlier draft and Paula Thomas for her assistance in locating references.

References

Adamson, D.P. (1971). Differentiated multi-track grouping vs. uni-track educational grouping in mathematics. Unpublished doctoral dissertation, Brigham Young University.

Alexander, K.L., & Cook, M.A. (1982). Curricula and coursework: A surprise ending to a familiar story. American Sociological Review, 47, 626-640.

Alexander, K.L., Cook, M.A., & McDill, E.L. (1978). Curriculum tracking and educational stratification. American Sociological Review, 43, 47-66.

Atkinson, J.W., & O'Connor, P. (1963). Effects of ability grouping in schools related to individual differences in achievement-related motivation (Final Report, Cooperative Research Project No. OE-2-10-024). Washington, DC: U.S. Dept. of Health, Education, and Welfare.

Bachman, A.M. (1968). Factors related to the achievement of junior high school students in mathematics. Unpublished doctoral dissertation, University of Oregon.

Bailey, H.P. (1968). A study of the effectiveness of ability grouping on success in first-year algebra. Dissertation Abstracts International, 28, 3061 A. (University Microfilms No. 68-1249).

Balow, I. (1964). The effects of "homogeneous" grouping in seventh grade arithmetic. Arithmetic Teacher, 12, 186-191.

Bangert-Drowns, R.L. (1986). Review of developments in meta-analytic method. Psychological Bulletin, 99, 388-399.

Barton, D.P. (1964). An evaluation of ability grouping in ninth grade English. Dissertation Abstracts International, 25, 1731 (University Microfilms No. 64-9939).

Bicak, L.J. (1962). Achievement in eighth grade science by heterogeneous and homogeneous classes. Unpublished doctoral dissertation, University of Minnesota.

Billett, R.O. (1928). A controlled experiment to determine the advantages of homogeneous grouping. Educational Research Bulletin, 7, 139.

Borg, W. (1965). Ability grouping in the public schools. The Journal of Experimental Education, 34, 1-97.

Borg, W. & Prpich, T. (1966). Grouping of slow learning high school pupils. Journal of Secondary Education, 41, 231-238.

Braddock, J.H. (1990). Tracking: Implications for student race-ethnic subgroups (Technical Report No. 1). Baltimore, MD: Center for Research on Effective Schooling for Disadvantaged Students.

Breidenstine, A.G. (1936). The educational achievement of pupils of differentiated and undifferentiated groups. Journal of Experimental Education, 5, 91-135.

Brophy, J.E., & Good, T.C. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 328-375). New York: Macmillan.

Campbell, A.L. (1965). A comparison of the effectiveness of two methods of class organization for the teaching of arithmetic in junior high school. Dissertation Abstracts International, 26, 813-814. (University Microfilms, No. 65-6726).

Chiotti, J.F. (1961). A progress comparison of ninth grade students in mathematics from three school districts in the state of Washington with varied methods of grouping. Unpublished doctoral dissertation, University of Northern Colorado, Greeley, CO.

Chismar, M.H. (1971). A study of the effectiveness of cross-level grouping of middle school underachievers for reading instruction. Unpublished doctoral dissertation, Kent State University.

Cochran, J. (1961). Grouping students in junior high school. Educational Leadership, 18, 414-419.

Coleman, J.S., & Karweit, N.L. (1972). Information systems and performance measures in schools. Englewood Cliffs, NJ: Educational Technology Publications.

Cottle, T.J. (1974). What tracking did to Ollie Taylor. Social Policy, 5, 21-24.

Crespo, M., & Michelna, J. (1981). Streaming, absenteeism, and dropping out. Canadian Journal of Education, 6, 40-55.

Dar, Y., & Resh, N. (1986). Classroom intellectual composition and academic achievement. American Educational Research Journal, 23, 357-374.

Drews, E.M. (1963). Student abilities, grouping patterns, and classroom interaction (Cooperative Research Project No. 608). Washington, DC: U.S. Department of Health, Education, and Welfare.

Esposito, D. (1973). Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments. Review of Educational Research, 43, 163-179.

Evertson, C.M. (1982). Differences in instructional activities in higher- and lower-achieving junior high English and math classes. Elementary School Journal, 82, 329-350.

Evertson, C., Sanford, J., & Emmer, E. (1981). Effects of class heterogeneity in junior high school. American Educational Research Journal, 18, 219-232.

Feldhusen, J.F. (1989). Synthesis of research on gifted youth. Educational Leadership, 46(6), 6-11.

Ferri, E. (1971). Streaming: Two years later. London: National Foundation for Educational Research in England and Wales.

Fick, W.W. (1963). The effectiveness of ability grouping in seventh grade core classes. Dissertation

Abstracts International, 23, 2753.

Findley, W.G., & Bryan, M. (1971). Ability grouping: 1970 status, impact, and alternatives. Athens: Center for Educational Improvement, University of Georgia. (ERIC Document Reproduction Service No. ED 060 595).

Fogelman, K., Essen, J. & Tibbenham, A. (1978). Ability grouping in secondary schools and attainment. Educational Studies, 4, 201-212.

Ford, S. (1974). Grouping in mathematics: Effects on achievement and learning environment. Unpublished doctoral dissertation, Yeshiva University.

Fowlkes, J.G. (1931). Homogeneous or heterogeneous grouping—which? The Nation's Schools, 8, 74-78.

Fox, L.H. (1979). Programs for the gifted and talented: An overview. In A.H. Passow (Ed.), The gifted and talented: Their education and development (pp. 104-126). Chicago: University of Chicago Press.

Gamoran, A. (1987). The stratification of high school learning opportunities. Sociology of Education, 60, 135-155.

Gamoran, A. (1989). Measuring curriculum differentiation. American Journal of Education, 97, 129-143.

Gamoran, A., & Berends, M. (1987). The effects of stratification in secondary schools: Synthesis of survey and ethnographic research. Review of Educational Research, 57, 415-435.

Gamoran, A., & Mare, R.D. (1989). Secondary school tracking and education inequality: Compensation, reinforcement, or neutrality? American Journal of Sociology, 94, 1146-1183.

Glass, G., McGaw, B., & Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

Good, T., & Marshall, S. (1984). Do students learn more in heterogeneous or homogeneous groups? In P. Peterson, L.C. Wilkinson, & M. Hallinan (Eds.), The social context of instruction: Group organization and group processes (pp. 15-38). New York: Academic Press.

Goodlad, J.I. (1983). A place called school. New York: McGraw-Hill.

Harrah, D.D. (1956). A study of the effectiveness of five kinds of grouping in the classroom. Dissertation Abstracts International, 17, 715. (University Microfilms No. 56-1114).

Heathers, G. (1969). Grouping. In R. Ebel (Ed.). Encyclopedia of educational research. (4th ed., pp. 559-570). New York: Macmillan.

Holy, T.C., & Sutton, D.H. (1930). Ability grouping in the ninth grade. Educational Research Bulletin, 9, 419-422.

Jones, J.D., Erickson, E.L., & Crowell, R. (1972). Increasing the gap between Whites and Blacks: Tracking as a contribution source. Education and Urban Society, 4, 339-349.

Kerckhoff, A.C. (1986). Effects of ability grouping in British secondary schools. American Sociological Review, 51, 842-858.

Kline, R.E. (1964). A longitudinal study of the effectiveness of the track plan in the secondary schools of a metropolitan community. Dissertation Abstracts International, 25, 324 (University Microfilms No. 64-4257).

Kulik, C.L., & Kulik, J. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. American Educational Research Journal, 19, 415-428.

Kulik, C., & Kulik, J. (1984). Effects of accelerated instruction on students. Review of Educational Research, 54, 409-425.

Kulik, J.A., and Kulik, C.L. (1987). Effects of ability grouping on student achievement. Equity and Excellence, 23, 22-30.

Lord, F.M. (1960). Large-sample convariance analysis when the control variable is fallible. Journal of the American Statistical Association, 55, 307-321.

Lovell, J.T. (1960). The Bay High School experiment. Educational Leadership, 17, 383-387.

Marascuilo, L.A., & McSweeney, M. (1972). Tracking and minority student attitudes and performance. Urban Education, 6, 303-319.

Martin, W.B. (1959). Effects of ability grouping on junior high school achievement. Dissertation Abstracts International, 20, 2810. (University Microfilms No. 59-1108).

Martin, W.H. (1927). The results of homogeneous grouping in the junior high school. Unpublished doctoral dissertation, Yale University.

McPartland, J.M., Coldiron, J.R., Braddock, J.H. (1987). School structures and classroom practices in elementary, middle, and secondary schools. Baltimore, MD: Johns Hopkins University, Center for Research on Elementary and Middle Schools.

Mikkelson, J.E. (1962). An experimental study of selective grouping and acceleration in junior high school mathematics. Unpublished doctoral dissertation, University of Minnesota.

Millman, J., & Johnson, M. (1964). Relation of section variance to achievement gains in English and mathematics in grades 7 and 8. American Educational Research Journal, 1, 47-51.

Newbold, D. (1977). Ability-grouping, the Banbury enquiry. Slough, England: National Foundation for Educational Research in England and Wales.

Newmann, F.M., & Thompson, J. (1987). Effects of cooperative learning on achievement in secondary schools: A summary of research. Madison: University of Wisconsin, National Center on Effective Secondary Schools.

Noland, T.K. (1985). The effects of ability grouping: A meta-analysis of research findings. Unpublished doctoral dissertation, University of Colorado.

Oakes, J. (1982). The reproduction of inequity: The content of secondary school tracking. Urban Review, 14, 107-120.

Oakes, J. (1985). Keeping track: How schools structure inequality. New Haven, CT: Yale University Press.

Persell, C.H. (1977). Education and inequality: A theoretical and empirical synthesis. New York: Free Press.

Peterson, R.L. (1966). An experimental study of the effects of ability grouping in grades 7 and 8. Unpublished doctoral dissertation, University of Minnesota.

Platz, E.F. (1965). The effectiveness of ability grouping in general science classes. Dissertation Abstracts International, 26, 1459-1460A (University Microfilms No. 65-6914).

Postlethwaite, K., & Denton, C. (1978). Streams for the future? Slogh, England: National Foundation for Educational Research.

Purdom, T.L. (1929). The value of homogeneous grouping. Baltimore: Warwick & York.

Reichardt, C.S. (1979). The statistical analysis of data from nonequivalent group designs. In T.C. Cook & D.T. Campbell, Quasi-experimentation: Design and analysis issues for field settings (pp. 147-205). Chicago: Rand McNally.

Rosenbaum, J.E. (1976). Making inequality: The hidden curriculum of high school tracking. New York: Wiley.

Rosenbaum, J.E. (1978). The structure of opportunity in school. Social Forces, 57, 236-256.

Rosenbaum, J.E. (1980). Social implications of educational grouping. Review of Research in Education, 8, 361-401.

Schafer, W. & Olexa, C. (1971). Tracking and opportunity: The locking-out process and beyond. Scranton, PA: Chandler.

Slavin, R.E. (1986). Best evidence synthesis: An alternative to meta-analytic and traditional reviews. Educational Researcher, 15 (9), 5-11.

Slavin, R.E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. Review of Educational Research, 57, 347-350.

Slavin, R.E. (1989). Achievement effects of substantial reductions in class size. In R.E. Slavin (Ed.), School and classroom organization (pp. 247-257). Hillsdale, NJ: Erlbaum.

Slavin, R.E. (1990). Cooperative learning: Theory, research, and practice. Englewood Cliffs, NJ: Prentice-Hall.

Slavin, R.E., Braddock, J.H., Hall, C., & Petza, R.J. (1989). Alternatives to ability grouping. Baltimore: Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.

Slavin, R.E., & Karweit, N.L. (1985). Effects of whole-class, ability grouped, and individualized instruction on mathematics achievement. American Educational Research Journal, 22, 351-367.

Slavin, R.E., Madden, N.A., & Leavey, M.B. (1984). Effects of team assisted individualization on mathematics achievement of academically handicapped and non-handicapped students. Journal of Educational Psychology, 76, 813-819.

Sorensen, A.B., & Hallinan, M.T. (1986). Effects of ability grouping on growth in academic achievement. American Educational Research Journal, 23, 543-557.

Stevens, R.J., Madden, N.A., Slavin, R.E., & Farnish, A.M. (1987). Cooperative integrated reading and composition: Two field experiments. Reading Research Quarterly, 22, 433-454.

Stoakes, D.W. (1964). An educational experiment with the homogeneous grouping of mentally advanced and slow learning students in the junior high school. Unpublished doctoral dissertation, University of Colorado, Boulder, CO.

Svensson, N.E. (1962). Ability grouping and scholastic achievement. Stockholm: Almqvist & Wirsell.

Thompson, G.W. (1974). The effects of ability grouping upon achievement in eleventh grade American history. Journal of Experimental Education, 42, 76-79.

Torgelson, J.W. (1963). A comparison of homogeneous and heterogeneous grouping for below-average junior high school students. Unpublished doctoral dissertation, University of Minnesota.

Trimble, K.D., & Sinclair, R.L. (1987). On the wrong track: Ability grouping and the threat to equity. Equity and Excellence, 23, 15-21.

Turney, A.H. (1931). The status of ability grouping. Educational Administration and Supervision, 17, 21-42, 110-127.

Vakos, H.N. (1969). The effect of part-time grouping on achievement in social studies. Dissertation Abstracts International, 30, 2271. (University Microfilms No. 69-20, 066).

Vanfossen, B.E., Jones, J.D., & Spade, J.Z. (1987). Curriculum tracking and status maintenance. Sociology of Education, 60, 104-122.

Wiatrowski, M., Hansell, S., Massey, C.R., & Wilson, D.L. (1982). Curriculum tracking and deliquency. American Sociological Review, 47, 151-160.

Wilcox, J. (1963). A search for the multiple effects of grouping upon the growth and behavior of junior high school pupils. Dissertation Abstracts International, 24, 205.

Willcutt, R.E. (1969). Ability grouping by content topics in junior high school mathematics. Journal of Educational Research, 63, 152-156.

Zweibelson, I., Bahnmuller, M., Lyman, L. (1965). Team teaching and flexible grouping in the junior high-school social studies. Journal of Experimental Education, 34, 20-32.

THE WHAT, WHY, AND HOW OF COOPERATIVE LEARNING

M. Lee Manning
Robert Lucking

The need for people to interact cooperatively and work toward group goals undoubtedly will increase during the 1990s. Yet, American education traditionally has emphasized individual competition and achievement, an approach that results in winners and losers and sometimes produces outright hostility among learners. Although research on competitive versus cooperative learning efforts can be traced to the early 1900s, considerable attention has focused on cooperative learning and its outcomes during the past two decades. In this article, we define cooperative learning, examine reasons for and benefits to students working in cooperative teams, and provide an overview of eight selected cooperative learning methods that hold potential for middle and secondary schools.

The current popularity and interest in cooperative learning, which undoubtedly will extend into the twenty-first century, must be viewed within a historical context. Cooperative learning methods have roots in social and psychological research and theory that focused on cooperation versus competition, some of which date back to the early 1990s (Slavin 1985). Maller (1929) focused his investigation on cooperation in the learning process, although research directed toward practical cooperative methods actually began in the early 1970s. During the 1980s, cooperative learning methods received considerable attention because of their potential to serve as an alternative to the traditional classroom that creates competition among learners rather than fostering a cooperative spirit (Slavin 1985).

What is Cooperative Learning?

Several scholars and proponents (Slavin 1983, 1985; Johnson and Johnson 1989/1990; Taylor 1989) have either offered definitions or provided the basic essentials of cooperative learning, and Kagen (1989) identified a broad repertoire of cooperative learning strategies. However, it was Slavin (1983) who defined the cooperative

*From **The Clearing House** (January/February 1991), Volume 64, Number 3, pp. 152-3. Reprinted with permission of the Helen Dwight Reid Educational Foundation. Published by Heldref Publications, 1319 18th Street, NW, Washington, DC 20036-1802.* *Copyright © 1991.*

learning process as a set of alternatives to traditional instruction systems, or, more specifically, techniques in which students work in heterogeneous groups of four to six members and earn recognition, rewards, and sometimes grades based on the academic performance of their groups. Such a definition does not imply, however, that all methods have the same processes and intended outcomes. Table 1 provides a brief overview of eight cooperative learning methods and shows how instructional efforts can share both similarities and dissimilarities. Commonalities may include cooperative efforts among learners, group goals, and the necessity for social skills (Johnson and Johnson 1989/1990; Slavin 1978; Aronson et al. 1978; Sharan and Sharan 1989/1990). Differences may include the fact that some methods are applicable to most grades and subjects (DeVries and Slavin 1978; Slavin 1978) and still other methods use individualization and are applicable only to mathematics in grades three through six (Slavin, Leavy, and Madden 1986).

Why Is There Current Interest in Cooperative Learning?

The current interest in cooperative learning stems from two broad forces: (1) the recognition that competitive educational environments encourage students to compete with one another rather than learn in a cooperative fashion and (2) evidence that suggests that cooperative learning, when properly implemented, has the potential for contributing positively to academic achievement, social skills, and self-esteem.

Traditionally, students have been taught in schools where competition reigned. Because learners competed with one another for grades and the teacher's attention and approval, students did not promote or encourage one another's academic efforts. Cooperative learning programs encourage students of all performance levels to work together in small groups toward group goals. Rather than competing for grades and other rewards as individuals, students help one another to learn and to achieve (Slavin 1987).

During the 1980s, considerable research (Slavin 1983, 1987; Johnson and Johnson 1987, 1989a) directed attention to the theory and practice of cooperative learning and promoted cooperative efforts over competitiveness. However, a few years earlier, Aronson et al. (1978) proposed a cooperative learning classroom that structured the learning process to make individual competitiveness and success compatible. In fact, success could occur only after the team had engaged in cooperative behavior. Instead of students' equating rewards and success with outperforming their competitors, learners achieved success as a consequence of listening to peers, helping and teaching each other, and helping others in the team to teach.

The second reason for increased interest in cooperative learning has to do with research indicating increases in academic achievement, enhancement of social skills, and the improvement of self-esteem. At the outset, it is imperative to warn against the fallacy of concluding that all cooperative learning methods are equally effective in making these improvements. Slavin (1988, 31) wrote, "I am becoming increasingly concerned about a widespread belief that *all* forms of cooperative learning are instructionally effective. This is emphatically not the case." Nevertheless, research studies have provided convincing evidence that cooperative learning, when properly implemented, can contribute positively to academic achievement, social skills, and self-esteem (see Johnson and Johnson [1989a] for a more comprehensive review).

Academic Achievement

Overall, most studies show that high, average, and low achievers gain equally from cooperative from cooperative learning experiences; however, some studies have shown greater gains for low achievers and others have shown the greatest gains for high achievers (Slavin 1985). Research studies on cooperative learning and academic achievement have, however, reached agreement on two aspects: (1) cooperative learning methods can and usually do have positive effects on student achievement, and (2) achievement effects do not result from all forms of cooperative learning (Slavin 1988, 1989/1990).

Cooperative learning methods that produce positive academic achievement share two features. First, they have group goals whereby team members work interdependently to earn teacher recognition or other forms of success. Second, they require individual accountability; that is, group success depends on individual contributions and learning of all members. For example, STAD, TGT, TAI, and CIRC (see Table 1 for explanations), which require group goals and individual accountability, consistently reveal positive results. The group investigation method (Sharan and Sharan 1989/1990) revealed positive effects on academic achievement even though it lacked group goals; its use of group rewards (similar to STAD, TGT, TAI, and CIRC) may serve a similar purpose (Slavin 1987).

Social Skills

It seems logical to assume that working and learning together as members of a team would contribute positively to students' social skills and to their overall ability to interact successfully with others. Research on social skills and gains either report positive effects in social areas or suggest the importance of teaching the social skills needed for successful small-group work.

Reporting the basic results of the Jigsaw classroom (see Table 1), Aronson et al. (1978) concluded that children grew to like their groupmates even more than others in the classroom. Similarly, Slavin (1987) found that students expressed greater liking for their classmates as a result of participating in cooperative learning methods.

Whether interpersonal and small-group skills are sufficiently important to warrant direct instruction has been an issue of recent debate. Johnson and Johnson (1989/1990) felt that simply placing students in groups and telling them to work together does not produce cooperation and high achievement. Certain social conditions increase students' efforts to achieve and to improve relationships with classmates: positive interdependence, face-to-face interaction, individual accountability, and social skills. For cooperative learning to be effective, students must get to know one another, communicate accurately and unambiguously, accept and support one another, and resolve conflicts constructively. Believing interpersonal and small-group skills to be as important as academic content, Johnson and Johnson (1989/1990) concluded that these skills should be taught just as systematically as a content subject.

Self-Esteem

Several research studies have concluded that cooperative learning teams increase students' self-esteem. This may be true for several reasons. First, feeling liked by peers and experiencing academic accomplishment are important components of students' self-esteem. Research has indicated that cooperative learning addresses both these

TABLE 1- Overview of Selected Cooperative Learning Methods

Method/Proponent	Brief Description/Comments
Learning Together (Johnson and Johnson 1987, 1989/90)	Emphasizing cooperative effort, Learning Together has five basic elements: positive interdependence (students beleive they are responsible for both their learning and the team's); face-to-face interaction (students explain their learning and help others with assignments); individual accountability (students demonstrate mastery of material); social skills (students communicate effectively, build and maintain trust, and resolve conflicts); group processing (groups periodically assess their progress and how to improve effectiveness). Uses four- or five-member heterogeneous teams.
Student Teams-Achievement Divisions (STAD) (Slavin 1978)	Four student learning teams (mixed in performance levels, sex, and ethnicity); teacher presents lesson, students work in teams and help others master material. Students then take quizzes; cooperative efforts are not allowed on quizzes; team rewards are earned. Applicable to most grades/subjects.
Teams-Games-Tournament (TGT) (Devries and Slavin 1978)	Using the same teacher presentation and teamwork as STAD, TGT replaces the quizzes with weekly tournaments in which students compete with members of other teams to contribute points to teams scores. Competition occurs at "tournament tables" against others with similar academic records. The winner of each tournament brings sic points to her or his team. Low achievers compete with low achievers (a similar arrangement exists for high achievers), which provides all students with equal opportunity for success. As with STAD, team rewards are earned. Applicable to most grades and subjects.
Jigsaw (Aronson, Blaney, Stephan, Sikes, and Snapp 1978)	Students are assigned to six-member teams to work on academic material that has been divided into sections. Each member reads a section; then, members of different teams meet to become experts. Students return to groups and teach other members about their sections. Students must listen to their teammates to learn other sections.
Jigsaw 2 (Slavin 1987)	Students work in four- or five-member teams as in TGT or STAD. Rather than being assigned specific parts, students read a common narrative (e.g., a chapter). Students also receive a topic on which to become an expert. Learners with the same topics meet together as in Jigsaw, and then they teach the material to their original group. Students take individual quizzes.
Team Assisted Individualization (TAI) (Slavin, Leavey, and Madden 1986)	Uses four-member mixed-ability groups (as with STAD and TGT); differs from STAD and TGT in that it combines cooperative learning and individualized instruction and is applicable only to mathematics in grades three through six. Learners take a placement test, then proceed at their own pace. Team members check one another's work and help with problems. Without help, students take unit tests that are scored by student monitors. Each week, the teacher evaluates and gives team rewards.
Cooperative Integrated Reading and Composition (CIRC) (Madden, Slavin, and Stevens 1986)	Designed to teach reading and writing in upper elementary grades, CIRC assigns students to different reading teams. Teacher works with one team, while other teams engage in cognitive activities: reading, predicting story endings, summarizing stories, writing responses, practicing decoding, and learning vocabulary. Teams follow sequence of teacher instruction, team practice, team preassessments, and quizzes. Quizzes may not be taken until the team feels each student is ready. Team rewards are given.
Group Investigation (Sharan and Sharan 1989/1990)	Groups are formed according to common interest in a topic. Students plan research, divide learning assignments among members, synthesize/summarize findings, and present the findings to the entire class.

components. Cooperative learning may lead to increased self-esteem because students have more friends and have more positive feelings about their school experiences and academic achievement (Slavin 1983, 1987, 1989/1990). Second, Aronson et al. (1978) concluded that there was a greater increase in self-esteem among children in the Jigsaw classroom than among the children in competitive classrooms. Third, in typical competitive classrooms, students become either winners or losers, with the latter developing negative attitudes toward themselves and school. With cooperative learning, students see each other as academically and socially competent colleagues rather than competitors. Such improved recognition of peers promotes better interpersonal relationships and leads in turn to increased self-esteem (Towson 1985).

How Can Educators Implement Cooperative Learning?

Implementing cooperative learning requires an understanding of the various types of cooperative learning, of the teacher's role in the process, and of the two previously noted essentials, that is, group goals and individual accountability.

Table 1 contains descriptions of eight cooperative learning methods and gives references for each. These eight methods are the most researched and, it is widely agreed, the most workable and effective methods. Student Teams-Achievement Divisions (STAD), Teams-Games-Tournament (TGT), Team Assisted Individualization (TAI), Cooperative Integrated Reading and Composition (CIRC), and Jigsaw 2 result from Slavin's (1987) work at the Johns Hopkins Center for Research on Elementary and Middle Schools.

Each of the various types of cooperative learning calls for a new role for the teacher. In the Jigsaw classroom (Aronson et al. 1978), the teacher, instead of being the major conveyer of knowledge, becomes the person who creates the process and environment that allow students to see each other as learning resources. Realistically, teachers have to abdicate some of their traditional power, such as being the sole dispenser of knowledge, grades, and discipline. Another change in roles includes viewing students as positive forces capable of learning on their own and from one another (Nickolai-Mays and Goetsch 1986).

As previously discussed, two essentials that significantly determine the degree of academic achievement are group goals and individual accountability (Slavin 1987). Other essentials include the five basic elements (see Table 1) proposed by Johnson and Johnson (1989b).

Effect of Cooperative Learning on Intergroup and Multiethnic Relations

People who help one another by joining forces to achieve a common goal generally feel more positively about each other and are willing to interact more positively when performing collective tasks. Rather than treating academic learning and social or intergroup relations as two distinct entities, cooperative learning contributes positively to overall intergroup relations and, particularly, to improving relations with multicultural/multiethnic students (Sharan 1985). Other research shows that students who learn cooperatively have greater liking for their cooperative learning classmates. Cooperative learning increases contact between students, provides a feeling of group membership, engages learners in pleasant activities, and requires that team members work toward a common goal (Slavin 1983).

Several research studies have indicated that cooperative learning programs can contribute positively to specific multiethnic populations. Aronson et al. (1978) concluded that both black and Anglo students in cooperative learning situations liked school better than did the same ethnic groups in competitive classrooms. Furthermore, for students of diverse populations, learning of this nature requires verbalization of ideas in a context of social interaction (Osterman 1990), and learning takes place through dialogue and conversation. Other research has suggested that working cooperatively seems to have particularly strong effects for Hispanic and black students, regardless of achievement levels (Slavin 1987). One study indicated that Anglo- and Asian-Americans in a learning situation similar to the Jigsaw method had more positive attitudes toward Mexican-Americans than did learners in competitive classes. A study of Jigsaw 2-related classes containing recent European and West Indian immigrants and Anglo-Canadians documented substantially more cross-ethnic friendships than in the control groups (Slavin 1983). Concerning specific cooperative learning methods, positive effects have been documented with STAD, TGT, TAI, and Jigsaw 2 (Slavin 1987). Studies of Group Investigation found that improved student attitudes and behaviors toward classmates of different ethnic backgrounds extended to classmates of different groups (Sharan, Amir, and Ben-Ari 1984). Cooperative learning arrangements also appear to facilitate relationships between handicapped and regular students (Yager, Johnson, Johnson, and Snider 1985). In conclusion, Slavin (1987) summarized that cooperative learning strategies apparently contribute to students' seeing each other in a positive light and forming friendships based on human qualities.

Summary

The advantages of learning cooperatively continue to become clear. First, social benefits accrue to students of heterogeneous ethnic and achievement backgrounds working together as functioning members of a team that focuses on common goals. Second, increases in academic achievement and self-esteem are also found among these students. Rather than pitting students against one another in competition for attention and grades, educators can select an appropriate cooperative learning method that effectively complements more traditional teaching styles and addresses their students' needs. Emphasizing cooperation over competition requires teachers to change their teaching roles from a power figure or conveyor of knowledge to a guide of learning events or to a resource person. The prospect of changing one's perception of what constitutes an effective teacher will undoubtedly prove unsettling to some educators. Nevertheless, the increasing need for all people to work together during the 1990s and beyond indicates that cooperative learning is an educational practice that contemporary educators must consider for their schools.

References

Aronson, E., N. Blaney, C. Stephan, J. Sikes, and M. Snapp. 1978. *The Jigsaw classroom*. Beverly Hills: Sage.

DeVries, D.L., and R.E. Slavin. 1978. Teams-Games-Tournament (TGT): Review of ten classroom experiments. *Journal of Research and Development in Education 12*: 28-38.

Johnson, D.W., and R. Johnson. 1987. *Learning together and alone: Cooperative, competitive, and individualistic learning*. Englewood Cliffs, N.J.: Prentice-Hall.

Johnson, D.W. and R. Johnson. 1989a. *Cooperation and competition: Theory and research*. Edina, Minn.: Interaction Book Company.

Johnson, D.W., and R. Johnson. 1989b. Toward a *cooperative* effort: A response to Slavin. *Educational Leadership 46*(7): 80-81.

Johnson, D.W., and R. Johnson. 1989/1990. Social skills for successful group work. *Educational Leadership 47*(4): 29-33.

Kagen, S. 1989. *Cooperative learning resources for teachers*. San Juan Capistrano, Calif.: Resources for Teachers.

Madden, N.A., R.E. Slavin, and R.J. Stevens. 1986. *Cooperative integrated reading and composition: Teacher's manual*. Baltimore: Johns Hopkins University, Center for Research on Elementary and Middle Schools.

Maller, J.B. 1929. *Cooperation and competition*. New York: Columbia Teachers College.

Nickolai-Mays, S., and K. Goetsch. 1986. Cooperative learning in the middle school. *Middle School Journal 18*: 28-29.

Osterman, K. 1990. Reflections practice: A new agenda for education. *Education and Urban Society* 22(2): 133-52.

Sharan, S. 1985. Cooperative learning and the multiethnic classroom. In *Learning to cooperate, cooperating to learn*, edited by R. Slavin, S. Sharan, S. Kagan, R. Lazarowitz, C. Webb, and R. Schmuck, 255-76. New York: Plenum.

Sharan, S., Y. Amir, and R. Ben-Ari. 1984. School desegregation: Some challenges ahead. In *School desegregation: Cross-cultural perspectives*, edited by Y. Amir and S. Sharan. Hillsdale, N.J.: Erlbaum.

Sharan, Y., and S. Sharan. 1989/1990. Group investigations expands cooperative learning. *Educational Leadership 47*(4): 17-21.

Slavin, R.E. 1978. Student teams and achievement divisions. *Journal of Research in Education* 12: 39-49.

———. 1983. *An introduction to cooperative learning*. New York: Longman.

———. 1985. An introduction to cooperative learning research. In *Learning to cooperate, cooperating to learn*, edited by R. Slavin, S. Sharan, S. Kagan, R. Lazarowitz, C. Webb, and R. Schmuck, 5-15. New York: Plenum.

———. 1987. *Cooperative learning*. 2nd ed. Washington, D.C.: National Education Association.

———. 1988. Cooperative learning and student achievement. *Educational Leadership* 47(4): 31-33.

———. 1989/1990. Research on cooperative learning: Consensus and controversy. *Educational Leadership 47*(4): 52-54.

Slavin, R.E., M.B. Leavy, and N.A. Madden. 1986. *Team accelerated instruction—Mathematics*. Watertown, Mass.: Mastery Education Corporation.

Taylor, R. 1989. The potential of small-group mathematics instruction in grades four through six. *Elementary School Journal 89*: 633-42.

Towson, S. 1985. Melting pot or mosaic: Cooperative education and interethnic relations. In *Learning to cooperate, cooperating to learn*, edited by R. Slavin, S. Sharan, S. Kagan, R. Lazarowitz, C. Webb, and R. Schmuck, 263-76. New York: Plenum.

Yager, S., R.T. Johnson, D.W. Johnson, and B. Snider. 1985. The effect of cooperative and individualistic learning experience on positive and negative cross-handicap relationships. *Contemporary Educational Psychology 10*: 127-38.

ALL STUDENTS CAN AND SHOULD LEARN SCIENCE---ENCOURAGING EQUITABLE CLASSROOMS

This section focuses on the accessibility of science and mathematics programs to students who typically are not proportionally represented in such programs. Low-income and minority students are the focus of Oakes' study, and women students are the focus of Kahle's article. Oakes finds that race, social class, and tracking limit the opportunities of students to achieve and participate in science and mathematics classes. She relies on data from the National Survey of Science and Mathematics Education (NSSME) conducted by the National Science Foundation. She specifically considers how student characteristics such as race, SES, and ability interrelate and also how these characteristics relate or correspond to differences in the quantity or quality of science and mathematics curricula.

Using a variety of research studies, Kahle examines the scarceness of women in science and proposes some methods for encouraging and retaining them in science programs. Her review includes studies on the elementary, secondary, and college/university levels, and she identifies the unique difficulties that women encounter on each level. Kahle then proposes some solutions based on the results of intervention programs currently being implemented and studied internationally. Some of these solutions include incorporating more spatial-visual activities, testing frequently, and designing curricula that appeal particularly to girls and that feature contributions of female scientists. Kahle's article goes beyond simply discussing the problem of recruiting and retaining women students in science to consider equitable science education for both women and men.

MULTIPLYING INEQUALITIES

Jeannie Oakes

I. The Effects of Student Characteristics on Opportunity

Relating students' race, social class, neighborhood, and career goals to resource allocation and instructional conditions is complicated, because while these classifications are individually important, they are also inseparable—no student has only one of them. In this section, therefore, we first consider the interrelatedness of these characteristics. We then use data from the NSSME to argue that the judgments schools make about students' ability should also be considered an equally inseparable characteristic that influences students' access to educational opportunities.

The Inseparability of Student Characteristics

Every student is a member of a racial group, a social class, and a type of community, and each can be identified as having a particular ability status (e.g., low, average, or high). In the United States, these various memberships and categories cluster.

African-American and Hispanic students[1] tend to be clustered in inner-city schools. Low-income[2] minority students are also concentrated in other urban and rural schools, but they are not often in the majority there.[3] On the other hand, middle- and upper-income white students are concentrated in and form a majority in suburban schools.[4] Most rural schools tend to be populated by white students, most of whom are neither very poor nor very well off. Nevertheless, as Table 2.1 shows, there are exceptions. Noticeably absent from the exceptions are all-minority suburban schools and predominantly minority high-wealth schools. Moreover, disproportionately more African-American and Hispanic minority students, poor students, and inner-city-school students are classified by schools as being low in academic ability and not likely to attend college. (We shall look more carefully at these relationships later in this section.)

Not only do these patterns exist, the clustering of particular categories is not serendipitous. Underlying social, political, and economic conditions provide substantive explanations for the fact that so many African-American and Hispanic youngsters are low-income, are living in inner-city neighborhoods, and are perceived to be less able

---

This selection first appeared as Chapters II and III in Jeannie Oakes (1990), ***Multiplying Inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science****, pp. 13-45. Santa Monica, CA: Rand Corporation. Rand publication number R-3928-NSF. Reprinted by permission.*

Table 2.1

SCHOOLS IN VARIOUS RACE, SES,[a] AND LOCALE CATEGORIES

(Percentage of all schools; N = 977)

	Inner City				Other Urban				Suburban				Rural				
Racial Composition	High Pov.	Mod. Pov.	Low Pov.	High Wlth.	High Pov.	Mod. Pov.	Low Pov.	High Wlth.	High Pov.	Mod. Pov.	Low Pov.	High Wlth.	High Pov.	Mod. Pov.	Low Pov.	High Wlth.	Total
0–10% white																	
Elementary	1.3	0.1	0.1	0	0.2	0.1	0.1	0	0	0	0	0	0.2	0.1	0	0	2.3
Secondary	1.3	0.5	0.1	0	0.4	0	0.3	0.1	0	0	0	0	0.4	0	0	0	3.2
10–50% white																	
Elementary	1.2	0.1	0.1	0	0.9	1.0	0.2	0	0.4	0.3	0.3	0	0.3	0.2	0.1	0	5.2
Secondary	0.4	0.2	0.1	0	0.8	1.8	0.8	0.1	0.3	0.5	0.2	0	0.2	0.7	0.3	0	6.6
50–90% white																	
Elementary	1.0	0.2	0.2	0	0.7	1.4	0.9	0.4	0.1	1.3	1.6	1.2	0.7	0.9	0.3	0.2	11.4
Secondary	0.6	0.9	0.6	0.1	0.5	2.3	2.2	1.2	0.5	1.9	2.2	2.5	0.8	3.2	1.5	0.3	21.3
90–100% white																	
Elementary	0.1	0	0.5	0.2	0.4	0.7	0.4	1.0	0.1	1.4	2.3	2.6	1.0	2.4	2.5	0.7	16.3
Secondary	0.2	0.1	0.5	0.1	0.3	1.2	1.9	1.4	0.3	2.2	4.3	7.4	0.8	5.8	6.1	1.1	33.9
Total	6.2	2.2	2.3	0.4	4.3	8.6	6.9	4.3	1.7	7.7	10.9	13.6	4.5	13.3	10.9	2.4	100

NOTE: Totals may not sum correctly due to rounding.

SOURCE: 1985-1986 NSSME.

[a] High poverty = more than 30 percent of students in school have parents who are either unemployed or on welfare; moderate poverty = 10 to 30 percent have parents who are unemployed or on welfare; low poverty and low-to-moderate wealth = less than 10 percent have parents who are unemployed or on welfare, and no more than 30 percent have parents in professional or managerial occupations; low poverty and high wealth = less than 10 percent have parents who are unemployed or on welfare, and more than 30 percent have parents in professional or managerial occupations.

than students from other backgrounds. These same dynamics also help to explain why such a large proportion of whites are economically better off, live in more affluent urban and suburban communities, and are more likely to be considered academically able.

While it has been argued that researchers should treat students' characteristics as inseparable, interactive influences on educational and occupational attainments (Grant and Sleeter, 1986), little is known about the simultaneous or combined effects of these characteristics. For example, while there are extensive data on the experiences of minorities *per se*, data for minority subgroups are rarely collected or analyzed separately. There is little data, for example, to document SES [socioeconomic status] differences within minority groups, such as those between low-income and middle-class African-American students.

The Relative Importance of Race and SES

One of the most puzzling questions concerning student characteristics is whether race or SES has the greater effect on children's opportunities for achievement and participation. Because these characteristics are nearly always studied separately, there is scant evidence with which to address this question. Race has received the bulk of the attention, since racial discrimination has been more often the subject of court actions and federal programs. The experiences of poor children are often extrapolated from studies of racial differences, since many of the poorest children in the United States are African-American and Hispanic, and most of the affluent children are white. But inferring the status of poor children in this country from the circumstances of African-American and Hispanic children grossly and stereotypically oversimplifies matters. Race and class probably cannot be effectively disentangled in attempts to understand the lower rates of achievement and participation of African-American and Hispanic students, but, as Table 2.1 illustrates, the overlap is not perfect. A significant disadvantaged sector of the school population—children who are poor and white—is often overlooked, as are middle-class minority students.

The importance of social class status in itself has been demonstrated in recent analyses. High School and Beyond (HSB) data illustrate that with other school and student factors (including race and ethnicity) controlled, students' SES accounts substantially for differences in mathematics achievement (Rock, Braun, and Rosenbaum, 1985). Similar relationships are found in Scholastic Aptitude Test (SAT) scores (College Board, 1985).

These findings compel us not only to consider race and social class separately, but also to look at how combinations of student characteristics relate to schooling experiences. The objective should not be to determine which single characteristic is the *most* important; analyses that control for some characteristics in the attempt to determine the effects of any single characteristic may mask the fact that all of them are important. We need to know how race, social class, and neighborhood *independently* affect students' learning opportunities. But we must also remember the overlaps among these characteristics, and whenever possible, we must explore how *combinations* of characteristics affect students' chances to learn.

The Relationships Between Ability Judgments and Opportunity

An often neglected key to understanding the distribution of critical features of science and mathematics education is the close connection between educators' judgments about students' intellectual ability and the educational experiences that follow from those judgments. The NSB has suggested that inequalities may stem from the "failure to recognize and develop talent" and "the erroneous belief that many students [lack] the ability to learn mathematics and science" (NSB, 1983:13).

Growing bodies of evidence show that assessments of students' intellectual abilities play a major role in the differential allocation of school experiences (Gamoran, 1987; Guthrie and Leventhal, 1985; Lee, 1986; Oakes, 1983, 1985, 1987). Judgments about academic ability often lead to the segregation of students into separate elementary and middle-school classes and to enrollment in different senior high school courses. These placements, in turn, can mediate students' opportunities to learn. In elementary and middle schools, students who appear to be slow are often placed in lower-level groups or remedial programs; students who seem to learn more easily are placed in high-ability groups. Research evidence suggests that these placements influence the pace and content of instruction and contribute to achievement disparities (Barr and Dreeben, 1983; Hallinan and Sorensen, 1983). At the senior high school level, judgments about students' ability influence decisions about curriculum track enrollment—whether students take college-preparatory, general, or vocational courses of study. Track enrollment, in turn, is critical in coursetaking and achievement (Lee, 1986; Lee and Bryk, 1988; Rock et al., 1984, 1985), as well as in curriculum content, instructional practices, and classroom learning environments (Oakes, 1985).

Class and track placements and subsequent differences in learning experiences have traditionally been explained as appropriate, given the apparent differences in students' performance at school. However, some evidence suggests that the ways elementary schools define ability and respond to students may help to solidify students' perceptions of their prospects for achievement and may eventually exaggerate initial performance differences among them (Rosenholtz and Simpson, 1984). Slavin (1987) shows that the popular practice of whole-class ability grouping at elementary schools is ineffective in increasing student learning. Similarly, secondary school tracking does not increase schools' overall achievement. Some studies have found that tracking may raise achievement for high-track students slightly, but when these gains are found, students in low tracks suffer a corresponding loss. These findings raise the possibility that in trying to accommodate students' differences with different educational experiences, schools may actually limit most students' opportunities to learn.[5]

Judgments about students' ability and the corresponding placement in homogeneous ability groups can affect the quality of students' educational experiences and achievement; therefore, tracking must be considered an important mechanism for distributing science and mathematics opportunities.

Race, Social Class, and Ability Classifications

Assessments of ability and placements in different classes appear to be particularly relevant to the educational experiences of poor and minority students, since the assessments often parallel race and class differences (Persell, 1977; Oakes, 1985; 1987; Rosenbaum, 1980). Low-income and minority elementary and middle-school students are more likely to experience initial learning difficulties; as a result, they are more likely to be judged as "low-ability" and placed in low-track and remedial classes or in special education programs (Persell, 1977; Rosenbaum, 1980; Slavin, 1987). Whites and upper-SES elementary students are more likely to be identified as able learners (and more likely to be considered "gifted and talented") and placed in enriched or accelerated programs (Darling-Hammond, 1985). In senior high school, African-Americans, Hispanics, and low-income students are enrolled more frequently in vocational and general programs, while whites and high-SES students are more frequently enrolled in academic programs (Rock et al., 1984, 1985).

There is a striking national pattern in the links between the tracking phenomenon and students' race and SES in the NSSME data, consistent with earlier research. Teachers at schools enrolling poor students and non-Asian minorities disproportionately judge their science and mathematics students to be of low ability.[6] In mixed schools, science and mathematics classes in which disproportionate numbers of minorities[7] are enrolled are highly likely to be perceived as low-ability classes.

Figure 2.1 shows that at both the elementary and secondary levels, schools in all four SES categories form classes that are homogeneous in ability, and that they do so to nearly the same extent.[8]

However, as Figs. 2.2 and 2.3 illustrate, despite the similarity in the percentages of homogeneous classes at both elementary and secondary schools, the *types of homogeneous groupings* differ quite dramatically in schools serving students of different SES levels.[9] Schools that have large concentrations of low-income students have a significantly greater percentage of low-ability groups. The percentage of low-ability classes drops and the percentage of high-ability groups increases significantly with higher school SES.

The minority enrollment at schools relates similarly to the patterns of ability groupings (Figs. 2.4, 2.5, 2.6). Figure 2.4 shows only small differences in the percentages of mixed or homogeneously grouped classes across types of schools. However, as Figs. 2.5 and 2.6 make clear, the types of ability-grouped classes differ significantly across schools with different racial compositions. Students who attend high-minority elementary and secondary schools are far more likely than students attending other schools to be enrolled in low-track science and mathematics classes and far less likely to be in high-track classes. Elementary school students' chances of being viewed as either average- or high-ability increase somewhat with the proportion of white students, the largest concentration of high-track classes occurring at the all-white schools. At secondary schools, the proportion of high-track classes increases dramatically as the proportion of white students increases, and there is a corresponding drop in the proportion of low-track classes.

There is also a second type of association between minority status and tracking. Tables 2.2 and 2.3 show the distribution of classroom ability classifications across classrooms with greater, roughly equal, and lower proportions of minority students than the minority enrollment at the school as a whole. We can consider these classes to be overrepresenting minorities, representative of the student population, or underrepresenting minorities.[10]

The distribution of ability-grouped classes at racially mixed elementary schools exhibited a highly significant pattern across the racial-composition categories.[11] Although disproportionately white classes were found to be about equally likely to be

87

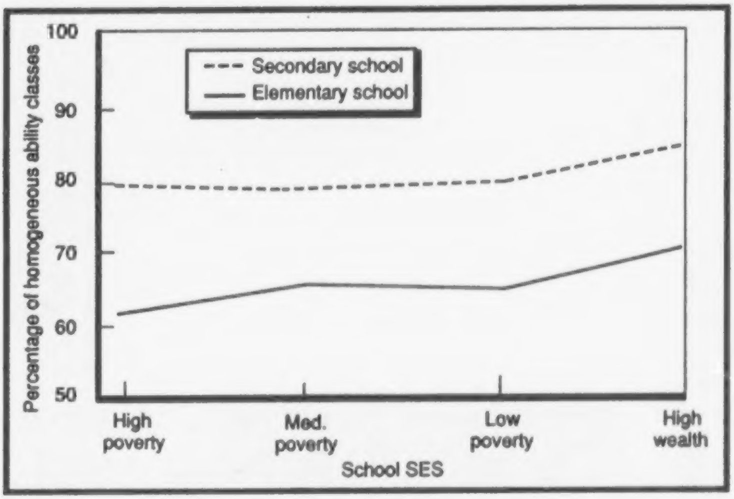


Fig. 2.1—Percentages of homogeneous ability classes, by school SES

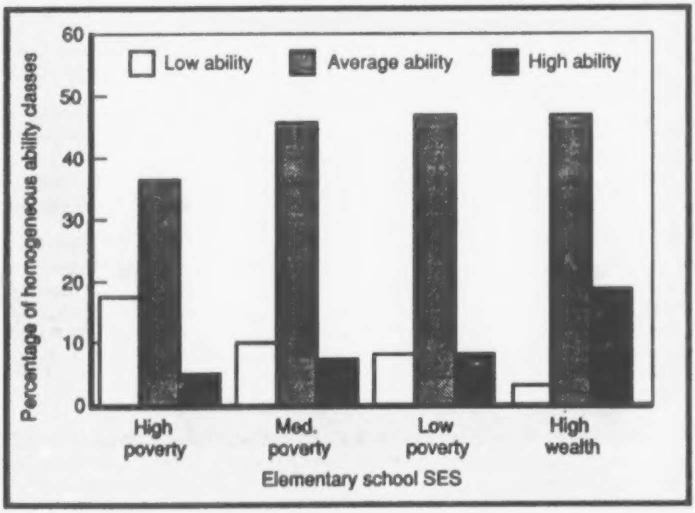


Fig. 2.2—Percentages of low-, average-, and high-ability classes in elementary schools, by school SES (F = 11.05, P < 0.001)

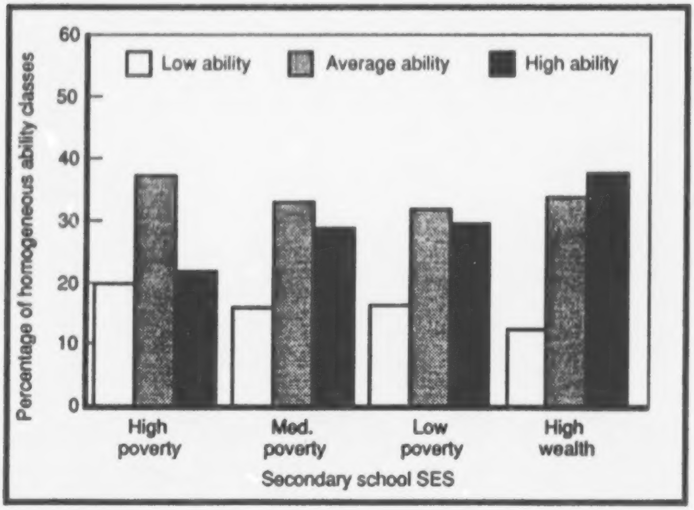


Fig. 2.3—Percentages of low-, average-, and high-ability classes in secondary schools, by school SES ($F = 3.92$, $P < 0.01$)

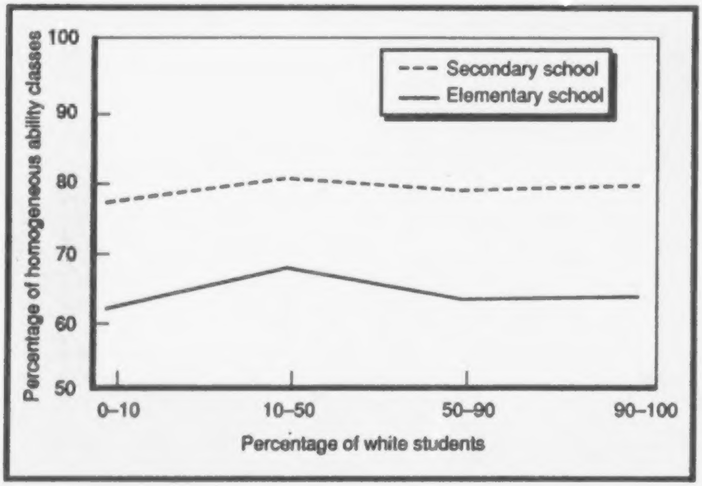


Fig. 2.4—Percentages of homogeneous-ability classes, by school racial composition

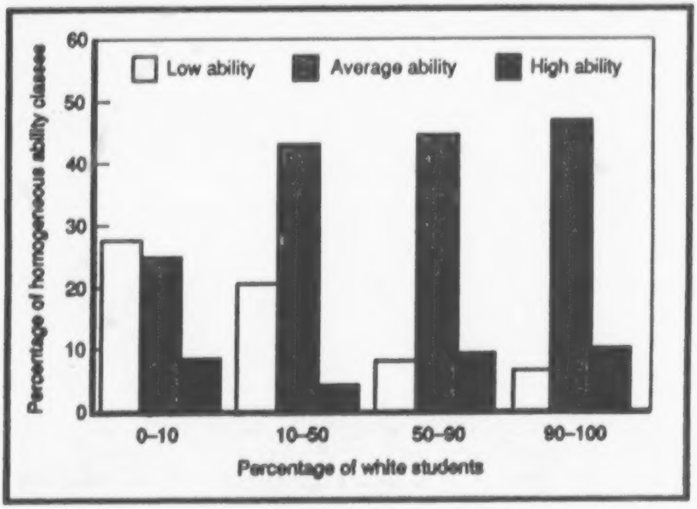


Fig. 2.5—Percentages of low-, average-, and high-ability classes in elementary schools, by school racial composition (F = 10.34, P < 0.001)

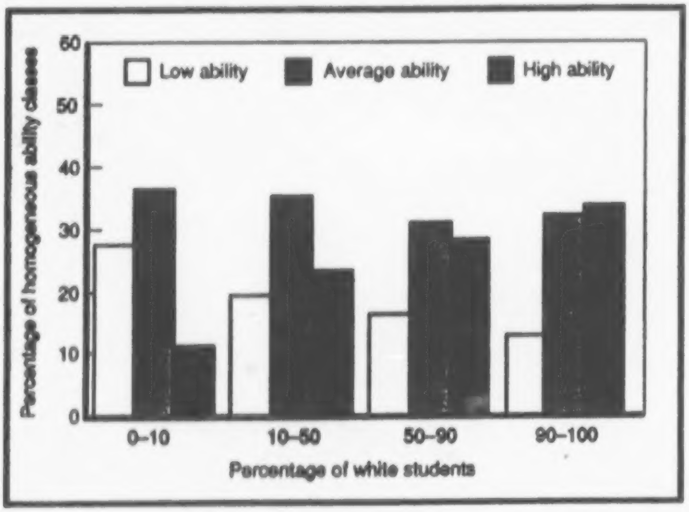


Fig. 2.6—Percentages of low-, average-, and high-ability classes in secondary schools, by school racial composition (F = 10.70, P < 0.001)

identified as low- or high-ability, disproportionately minority classes were *seven times more likely to be identified as low-ability than as high-ability.*

The association is even more dramatic in secondary schools: Not only were disproportionately minority secondary school classes far more likely to be judged low-ability (two-thirds of all such classes), disproportionately white classes were rarely judged this way. Differences in judgments about high ability were equally dramatic: More than

Table 2.2

ABILITY LEVELS OF CLASSES IN ELEMENTARY SCHOOLS, BY RACIAL COMPOSITION OF CLASS RELATIVE TO SCHOOL ENROLLMENT

(In percent)

Class Enrollment[a] Relative to School Enrollment	Class Ability Level		
	Low	Average	High
Fewer minorities	22	52	26
Same (within 10%)	14	74	12
More minorities	30	66	4

NOTE: Table includes classes at mixed-race elementary schools (10 to 90 percent minority) only.

[a] Effect of class racial composition (relative to school's) is significant at the 0.001 level (F = 8.54).

Table 2.3

ABILITY LEVELS OF CLASSES IN SECONDARY SCHOOLS, BY RACIAL COMPOSITION OF CLASS RELATIVE TO SCHOOL ENROLLMENT

(In percent)

Class Enrollment[a] Relative to School Enrollment	Class Ability Levels		
	Low	Average	High
Fewer minorities	5	38	57
Same (within 10%)	21	50	29
More minorities	66	25	9

NOTE: Table includes classes at mixed-race elementary schools (10 to 90 percent minority) only.

[a] Effect of class racial composition (relative to school's) is significant at the 0.001 level (F = 143.45).

half of the disproportionately white classes were judged high-ability, compared with fewer than 10 percent of the disproportionately minority classes.

Our analyses of schools with different racial compositions and of classes at racially mixed schools show that African-American and Hispanic minorities face two potential barriers to science and mathematics opportunities. First, their access to high-track science and mathematics classes diminishes as the minority enrollment at their schools increases. Second, those who attend racially mixed schools are more likely than their white peers to be placed in low-track classes. There is thus a "double jeopardy" effect for African-American or Hispanic students. These relationships suggest strongly that minority students will suffer whatever disadvantages accrue to students in low-ability classes.[12]

Tracking is commonly viewed as a neutral, educationally sound response to a wide range of student aptitude and achievements, but the evidence we offer here confirms earlier work showing that such groupings are easily and commonly confounded with race and social class. Moreover, much prior research suggests that the differences in opportunities provided to ability-grouped classes *limit* instruction, rather than fine-tune it in ways that accommodate individual differences and promote learning (Oakes, Gamoran, and Page, in press). We examined the distribution of opportunities by the track level of science and mathematics classes to further investigate the possibility that in their efforts to accommodate differences in ability with different educational experiences, schools actually limit some students' opportunities to learn.

Because assessments of low academic ability and placements in nonacademic programs occur more frequently among low-income and minority students, the combined effects of ability and other background characteristics are of special importance in attempts to uncover factors related to underachievement and low participation among these groups. The recurrence of this relationship at both the school and classroom levels underscores the importance of considering ability grouping in any study of the relationship of students' characteristics to the distribution of opportunity.

Notes

1. Our analyses of minority students focus on African-Americans and Hispanics. The small number of Native Americans in the sample does not permit us to extend our findings to that group, although we expect that their experiences parallel those of the African-Americans and Hispanics. We did not include the Asian students in our minority category, since our purpose was to describe the experiences of minority groups that are underrepresented in science and mathematics. We grouped Asian students with whites in our analyses; separate analyses excluding Asians did not yield different results, however, because of the small size of this group.

2. At the school level, we created variables for SES based on principals' reports of school characteristics. Because recent work has established its importance for a number of schooling outcomes (Orland, 1988), the concentration of poor students was used as the basis for our SES measure, with schools reporting low concentrations further differentiated by the proportion of students from wealthy backgrounds. We are well aware that determining school and student SES is always difficult with surveys, and completely satisfactory proxies are rarely found. Our questionnaires asked principals to list the percentage of students with parents in the following occupational categories: (1) professional and managerial; (2) sales, clerical, technical, or skilled workers; (3) factory or other blue collar workers; (4) farm workers; (5) persons not regularly employed; (6) persons on welfare. We felt that while principals were likely to be fairly accurate in estimating percentages of students from families at the high and low ends of this occupational

scale, they would be less able to discriminate among those parents who held mid-level jobs (categories 2 and 3). Moreover, we considered "farm workers" an ambiguous category, which principals might interpret to include a wide range of occupational levels, from poor, itinerant field hands to owner-operators. We therefore constructed our SES categories from responses to the extreme categories (1, 5, and 6), devising measures that reflect the extent to which wealth and poverty are represented. The following SES categories were defined: (1) high poverty (more than 30 percent of the students have parents who are either unemployed or on welfare); (2) moderate poverty (between 10 and 30 percent of the students have parents who are either unemployed or on welfare); (3) low poverty/low to moderate wealth (less than 10 percent of the students have parents who are either unemployed or on welfare, and no more than 30 percent have parents in professional or managerial occupations); (4) low poverty/high wealth (less than 10 percent of the students have parents who are either unemployed or on welfare, and more than 30 percent of the students have parents in professional or managerial occupations).

3. Categories of schools that differed in racial composition were defined as follows: (1) minority—less than 10 percent of the students are white; (2) mixed (mostly minority)—10 to 50 percent are white; (3) mixed (mostly white)—50-90 percent are white; (4) white—more than 90 percent are white.

4. We did not define new variables for determining a school's community type. Principals were asked to choose whether their school was best described as (a) inner city, (b) urban, but not inner city, (c) suburban, or (d) rural.

5. See Oakes, Gamoran, and Page (in press) for a review of this literature.

6. The ability or track level of a class represents the teachers' categorization of the class as including (1) a wide range of ability levels, (2) predominantly low-ability students, (3) predominantly average-ability students, or (4) predominantly high-ability students. We must emphasize two limitations of analyses of ability groupings in the NSSME data. First, we are dealing with perceived ability as a *characteristic of the class*—we can legitimately talk about opportunities available to students in low-track classes, etc., but we can say nothing about individual ability. Second, we are dealing with *teachers' perceptions* of ability—we cannot assume that ability means the same thing from teacher to teacher, or from school to school. Finally, while it is impossible to determine from the NSSME questionnaire whether a formal tracking system is in place at a school, the percentages of homogeneous groups reported by teachers agree very closely with the percentages of elementary schools reporting that they use whole-class ability grouping (Slavin et al., 1989).

7. We used a variable created from principals' descriptions of school characteristics and teachers' descriptions of sampled classes to show the overrepresentation or underrepresentation of African-American and Hispanic students. We developed three categories for this variable: (1) disproportionately white—smaller percentage of minority students in the classroom than in the school (differences can range from 10 percent to a theoretical 100 percent); (2) proportionate—same percentage as in the school (within 10 percent in either direction); and (3) disproportionately minority—larger percentage of minority students than in the school (differences can range from 10 percent to 100 percent).

8. The most-affluent schools seem to group students homogeneously more frequently than the least-affluent schools, although the 5 to 10 percent difference is not statistically significant.

9. Percentages in Figs. 2.2 through 2.6 do not sum to 100 because mixed-ability classes are excluded from these charts.

10. For the school-level analysis, we used all secondary and middle schools in the data set. But because the preponderance of 90 percent or more white schools in the sample, the classroom level analysis was conducted for classes in schools with mixed racial composition (i.e., categories 2 and 3 of the school percentage-white variable). This analysis allows us to control for the racial composition of the schools when we consider the relationship between minority

enrollment and ability classification at the classroom level.

11. The tests of significance were conducted using class weights. For the school-level analysis, teachers' responses were averaged by school, and their weights were summed for each school, yielding a single observation. However, the data displayed in table form represent the teachers' responses directly.

12. Because of the overlaps between minority status and poverty, a similar double disadvantage probably exists for low-income students. We were unable to test this hypothesis, however, because the NSSME survey collected no SES data at the classroom level.

II. Access to Science and Mathematics Programs

Obviously, the mathematics and science curriculum that is taught makes a difference in what students learn: When teachers teach particular mathematics and science topics, concepts, processes, and skills, students are more likely to learn them (Crosswhite et al., 1985; Husen, 1967; McKnight et al., 1987; Wolf, 1977). In this section, we examine the effects of school and classroom characteristics on the mathematics and science knowledge various groups of students have an opportunity to learn.

At the elementary level, the most basic—albeit limited—indicator of students' access to science and mathematics knowledge is the amount of time teachers spend teaching these subjects. An equivalent indicator at the secondary level is the number of science and mathematics courses schools offer. But to really understand the type of knowledge that students have an opportunity to learn, we also need to know the curricular goals that guide science and mathematics instruction for different students; the topics and skills being taught and the depth of course coverage; the numbers of topics covered; the scientific accuracy of the content of science and mathematics lessons; and the suitability of the lessons to students' developing cognitive abilities. There is little reliable data available, since large-scale research on these issues is difficult and costly.

In this section, we present new evidence about the quantity and quality of the science and mathematics curriculum that different groups of students experience in schools and classrooms and place this evidence in the context of findings from other research. First, we look at the extent of programs, measured by the amount of time spent on science and mathematics in elementary schools and *how many courses* in these subjects are offered in secondary schools. Then we consider the depth and rigor of the content in secondary school programs, as measured by the types of courses schools offer. We attempt to show the similarities and the differences in students' access to science and mathematics knowledge, and we examine the relationship of students' race, social class, community, and ability status to differences in their curricula.

Time Spent on Science and Mathematics in Elementary Schools

As noted above, the amount of class time that elementary teachers devote to teaching science and mathematics is the most basic indicator of students' access to these subjects, but it indicates little about the time students spend engaged in lessons or about the pace, content, and quality of instruction. Nevertheless, there is reason to believe that the amount of time spent studying a subject does influence the amount of learning that takes place.[1] This relationship is strongest in mathematics and science, since these subjects, unlike reading and social studies, are rarely learned in informal settings (Husen, 1967).

The time devoted to various subject areas varies considerably among schools. Goodlad (1984) surveyed teachers in thirteen elementary schools and found that their reported time spent teaching mathematics ranged from 45 to 66 minutes per day. On average, mathematics occupied 20 percent of the total instructional time in these schools. Science typically occupied less time, but the variation among the schools was greater, with reported time ranging from 16 to 64 minutes. The average time devoted to science instruction was 28 minutes, about 10 percent of the total instructional time in these schools.

The NSSME data yield similar findings. Teachers of self-contained classes in grades K-3 reported spending an average of 43 minutes per day on mathematics instruction and 18 minutes on science.[2] Their counterparts teaching grades 4-6 reported spending 52 minutes on mathematics and 29 minutes on science. At both levels, the variations in science were greater than those in mathematics. Schools using "specialists" to teach science and/or mathematics typically devoted more time to science and slightly less time to mathematics (Weiss, 1987).

However, the variation in mathematics instructional time was not random. Schools and classrooms serving different groups of students spent significantly different amounts of time on mathematics lessons. Figure 3.1 shows that the number of minutes per day spent in mathematics instruction at the NSSME elementary schools differed by the percentage of low-income students enrolled.

Slightly more time per day was spent on mathematics at schools enrolling large concentrations of low-income students than at other types of schools: 50 minutes, on average, compared with 47, 43, and 44 minutes at schools serving increasingly advantaged populations. Similar time differences occurred among schools serving different racial groups and in different types of communities. Schools with the largest percentages of African-American and Hispanic students and those in inner cities

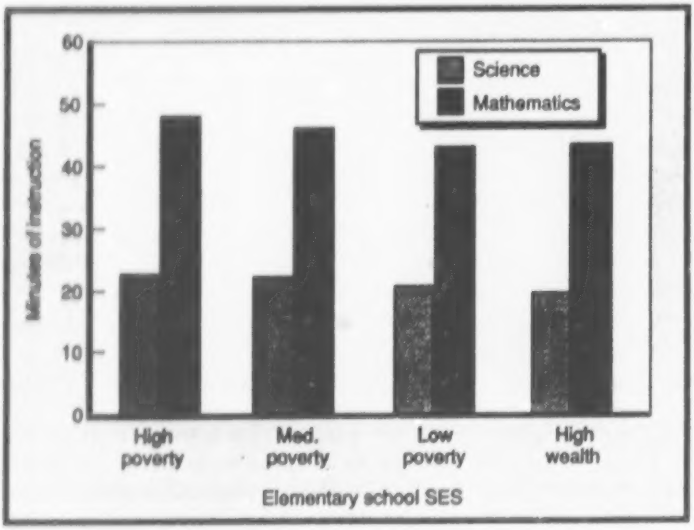


Fig. 3.1—Time spent on science and mathematics in elementary schools serving different student populations (for science, no significant differences; for math, F = 4.92, P < 0.01)

devoted the most time to mathematics. Interestingly, the small amount of additional time in mathematics in low-SES, high-minority, and inner-city schools did not seem to be gained at the expense of science instruction, for which time allotments were similar across school types. Because these schools tend to have longer school days, they can accommodate longer periods of mathematics instruction. For example, the low-SES schools averaged 171 instructional minutes per day, and the highest-SES schools averaged 156 minutes. The percentage of the school day spent on mathematics was a relatively constant 28 percent across schools.

On its face, it is not surprising that teachers at low-SES, high-minority, and inner-city schools would spend more time on mathematics, since most of these schools qualify for federal assistance in the form of Chapter 1 funds targeted at improving disadvantaged students' basic skills (96 percent of the schools in our low-SES category reported that they qualified for these funds). However, federal assistance did not seem to affect directly the time spent on mathematics instruction. What it did affect was the time teachers said they spent on science. In all but the most affluent group of schools (only 15 percent of which qualified for federal funding), schools with Chapter 1 programs spent more time on science.[3] We speculate that raising students' mathematics achievement is thought to be important enough to command a large share of time *with or without Chapter 1 funding*. But the extra resources—in the form of specialist teachers and instructional materials—may enable participating schools to free up more teacher time and resources for science instruction.

In contrast to these school differences, the track level of elementary science and mathematics classes does not seem to affect the allocation of instructional time. Within NSSME schools of different types, students in low-, average-, and high-ability classes spent equivalent amounts of time on lessons. Because schools often regulate the minimum amount of class time for various subjects, teachers may have little discretion about how much time they direct students of different ability levels to spend on lessons.[4]

Thus, elementary schools with large concentrations of students who typically do poorly in mathematics seem to be attacking this problem by spending somewhat more time on mathematics lessons, presumably in hopes of giving disadvantaged and minority students a good start in mathematics. However, across all schools extra mathematics time is not being provided in those classes for the students who are perceived to be the least able.[5]

Our cross-sectional data do not permit us to determine whether the increased time allocations in schools with large concentrations of low-income and minority students are long-standing or reflect recent changes—perhaps in response to the press for additional time on basic subjects in recent school reform proposals (e.g., *A Nation at Risk*). Nor can these data indicate whether the extra mathematics time in disadvantaged and minority schools is spent in ways that are likely to help students overcome their historic patterns of low achievement. We know, for example, that the time teachers allocate to lessons is far less important than the time students are actually engaged in appropriate tasks (Berliner, 1979; Brophy and Good, 1986). Despite the limitations of these findings, however, achievement data showing a slow but steady decrease in the black-white mathematics gap in the elementary grades over the past decade suggest that some factor in basic skills instruction—perhaps even this small additional time allocation—may be having a modest benefit. However, other aspects of elementary school students' experiences, discussed in later sections, may offset this advantage.

Science and Mathematics Programs in Secondary Schools

Secondary students' access to mathematics and science can be judged roughly by the courses in these subjects that schools make available. In this section, we are interested in (1) the extent of science and mathematics programs, (2) the content and rigor of the courses that make up those programs, and (3) the extent to which schools offer typical "gatekeeping" courses, i.e., courses that are usually prerequisites for participation in higher-level science and mathematics.

The Extent of Programs

Secondary schools vary considerably in the time they allocate to science and mathematics. Goodlad (1984) found that an average of 17 percent of full-time teaching positions (equivalent to the percentage of instructional time) were devoted to mathematics in a sample of junior high schools. However, the range among the schools was 13 to 22 percent. In science, the average was 13 percent, but it ranged from 7 to 20 percent. Slightly less variability was found at the senior high school level, where, on average, 13 percent (with a range of 9 to 20 percent) of the teaching positions were allocated to mathematics, and 11 percent (with a range of 8 to 15 percent) were allocated to science.

What makes this variation worth noting is the connection between students' exposure to subjects and their achievement. Considerable evidence from national surveys (e.g., the National Assessment of Educational Progress (NAEP), National Longitudinal Study (NLS), and HSB) attests to the importance of high school coursetaking for mathematics and science achievement (Jones et al., 1986; Welch, Anderson, and Harris, 1982). Thus it is not surprising that the number of courses offered relates to what and how much students learn (Peng, Owings, and Fetters, 1981; Rock et al., 1985).

The extent of programs in secondary schools is a rough corollary of the time devoted to classes in elementary schools. Here, we are interested in the *numbers of class sections* offered in science and mathematics.[6] The numbers indicate *how much* science and mathematics is available and the extent to which students are taking advantage of available opportunities—either because of their own or their parents' wishes or because the school requires or presses students to enroll in science and mathematics courses. Of course, they reveal nothing about what *kinds* of courses are available. It is highly significant that differences in the extent of the programs at the NSSME secondary schools are not random, but rather are fairly consistently related to school population and community characteristics.[7]

Middle Schools and Junior High Schools. Figures 3.2 and 3.3 show the differences in the extent of the mathematics and science programs at the junior high schools in the NSSME sample by school SES and racial composition.[8]

These patterns are quite different from those related to time allocations in elementary schools. Students attending the junior high and middle schools with the largest concentrations of *low-income students* had access to considerably less-extensive

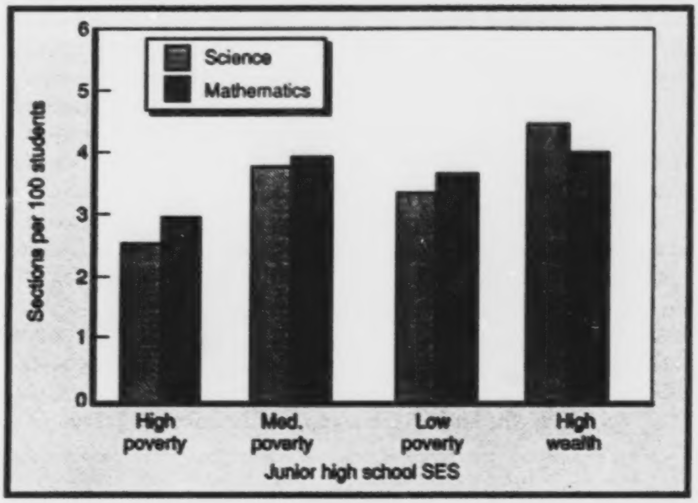


Fig. 3.2—Mathematics and science classes per 100 students in grade 6–9 junior high schools, by school SES (for mathematics, F = 4.39, P < 0.01; for science, F = 13.69, P < 0.001)

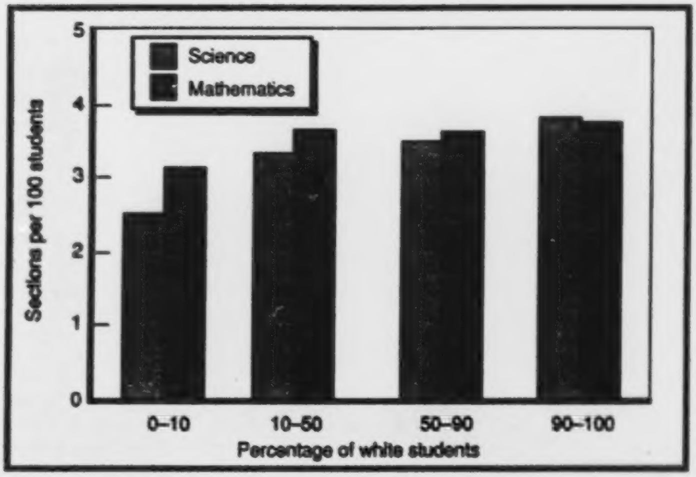


Fig. 3.3—Mathematics and science classes per 100 students in grade 6–9 junior high schools, by school racial composition (for mathematics, differences not significant; for science, F = 3.64, P < 0.05)

programs in both science (approximately 2.5 classes per 100 students) and mathematics (approximately 3 classes per 100 students) than those attending the most affluent schools (slightly more than 4 classes per 100 students in each subject); differences in science exceeded those in mathematics. The findings are only slightly different when the schools' racial makeup is considered. Science programs at schools whose student populations are more than 90 percent white were found to be significantly more extensive (nearly 4 classes per 100 students) than those at schools with similar percentages of minorities (about 2.5 classes); no significant race-related differences were found in the size of mathematics programs.

These findings, while perhaps disappointing, are not surprising. Many junior high school and middle school students are required to take only one semester or one year of science; additional science courses are either electives or recommended for high-achieving students. What the findings suggest is that high-SES and predominantly white schools either have greater science requirements or offer greater numbers of optional science courses.

The finding that low-SES junior high schools offer fewer sections of mathematics raises more troubling questions. In nearly every state, junior high school and middle school students are required to take mathematics each year, so mathematics classes in low-SES schools may be consistently larger than those in other types of schools—large enough to reduce the relative number of mathematics sections offered. Our data, in fact, do show such a trend. The lowest-SES schools had significantly larger mathematics classes than did the most affluent schools (averaging 26 and 23 students per class, respectively).[9] This means that the junior high schools serving the largest concentration of low-income students allocated less time and fewer teachers to mathematics, even though all students were probably enrolled. Community type had little effect on the extent of junior high school programs. This suggests that low-SES schools in both inner-city and rural areas (the communities where most low-SES schools are found) offer students fewer science and mathematics opportunities.

These findings suggest that the possible *advantages* of more time on mathematics instruction in elementary schools serving low-income and minority students may disappear when students reach middle school or junior high school.

Senior High Schools. Science and mathematics programs are similar in size at most senior high schools, but high-SES schools have significantly larger science programs than other types of schools, and rural schools have the least extensive mathematics programs.[10]

Content and Rigor of Programs

Differences in the number of class sections reflect only the most general differences among programs. Data about the content and level of courses (e.g., advanced, general, or remedial) add considerable information, since programs at schools offering the same number of class sections may differ significantly in depth and rigor. For example, if a school offers no biology courses, no one at that school will have a chance to learn biology. These differences are important because the particular courses students take affect their achievement (Jones, 1984; Jones et al., 1987; Peng, Owings, and Fetters, 1981; Sells, 1982).

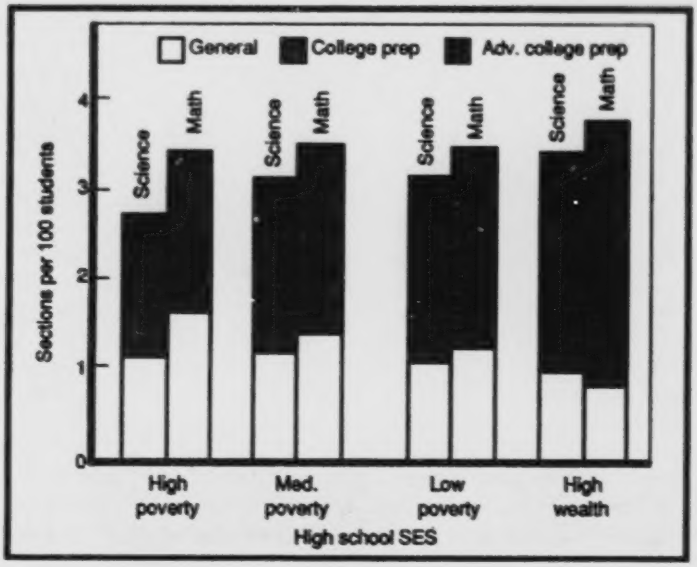


Fig. 3.4—Mathematics and science classes per 100 students in senior high schools, by school SES

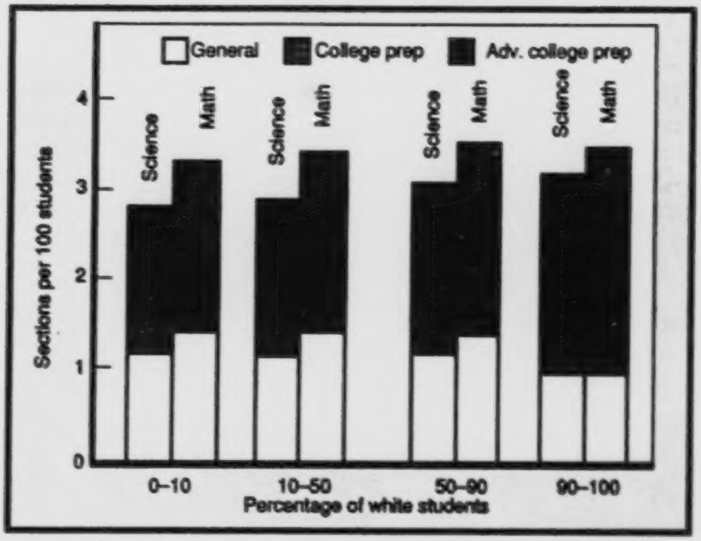


Fig. 3.5—Mathematics and science classes per 100 students in senior high schools, by school racial composition

Table 3.1

SIGNIFICANCE OF SES, RACE, AND LOCALE DIFFERENCES FOR SENIOR HIGH SCHOOL COURSE OFFERINGS

	Science		Mathematics	
Independent Variable	**F**	**P <**	**F**	**P <**
General Classes				
SES	1.30	0.27	12.42	0.001
% White	2.02	0.11	7.80	0.001
Locale	1.51	0.21	3.71	0.02
College Prep Classes				
SES	8.23	0.001	11.23	0.001
% White	5.48	0.001	3.00	0.03
Locale	3.45	0.02	7.32	0.001
Advanced College Prep Classes				
SES	11.30	0.001	16.54	0.001
% White	6.97	0.001	6.67	0.001
Locale	4.08	0.01	10.49	0.001
All Classes				
SES	3.27	0.02	2.06	0.11
% White	1.22	0.3	0.34	0.8
Locale	2.20	0.09	4.05	0.01

Because course offerings can constrain or encourage students' coursetaking in science and mathematics, differences in programs may contribute to the achievement gaps among different student groups. HSB data show that senior high schools serving predominantly poor students typically offer fewer advanced placement (AP) courses (Ekstrom, Goertz, and Rock, 1988) and enroll proportionately fewer students in those they offer (Jones, Burton, and Davenport, 1984), and that students in college-preparatory programs at low-SES schools (the schools most minorities attend) typically take fewer academic courses (including mathematics and science courses) than their college-bound peers attending more-advantaged schools (Rock et al., 1985). These lower rates of academic coursetaking are linked to African-American students' lower levels of achievement in these subjects (Jones, 1984; Pallas and Alexander, 1983).

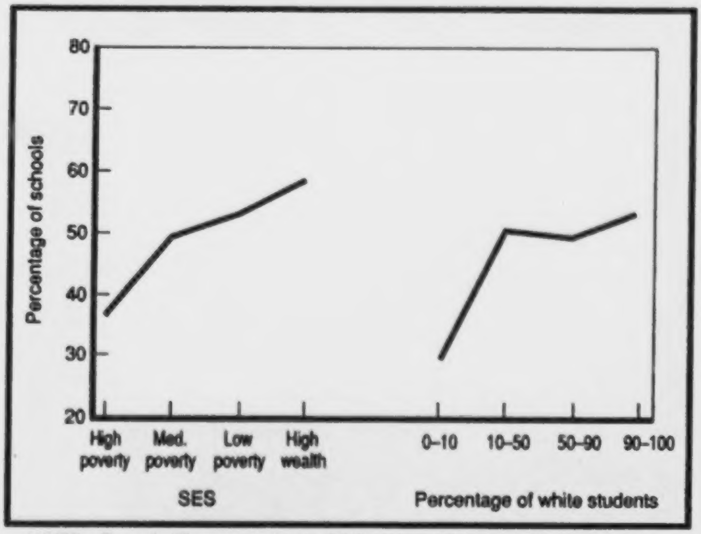
Figures 3.4 and 3.5 show the number of sections of general, college-preparatory, and advanced college-preparatory courses at the senior high schools in the sample.[11]

When the number of class sections is broken down by types of courses, the distribution patterns are significant or highly significant for all but general science classes. As the proportion of low-income and minority students at a school increases, the relative proportion of college-preparatory or advanced course sections decreases. Table 3.1 shows the significance of these differences.

The differences in the level of the classes offered cannot be accounted for by class size differences, as was the case among the junior highs. In fact, some of the differences in the sizes of various categories of classes in different types of schools compound the difference in opportunities among schools. For example, while no overall SES-related differences were found in the number of students in science classes, the size of college-preparatory and advanced college-preparatory mathematics classes increased with the affluence of the school.[12] This means that relatively more students were taking these courses at highly affluent schools than simple comparisons of the numbers of sections would indicate.

The Availability of Gatekeeping Courses

A third important dimension of students' access to science and mathematics knowledge is the extent of critical "gatekeeping" courses offered at their schools, that is, courses that are especially important in qualifying students for post-high-school work in science and mathematics. At the senior high school level, the most critical course is calculus, because it is a prerequisite for entry into most science-, mathematics-, and technology-related majors at college. Without high school calculus, students must take beginning calculus classes in college. In many cases, this can make obtaining a baccalaureate degree in a quantitative field in four years very difficult. At the junior high school level, eighth grade algebra and ninth grade geometry are critical gatekeepers, since students who take these courses early are on track by grade 12 without having to double up classes or take mathematics courses during the summer.[13]



NOTE: Overall differences due to SES and racial composition were not significant, although groups 1 and 4 differed with P-values approaching significance (P = 0.086 for SES, P = 0.08 for racial composition).

Fig. 3.6—Junior high schools offering accelerated mathematics classes, by school SES and racial composition

Junior High Schools. We performed parallel analyses for the availability of algebra for eighth graders and geometry for ninth graders in the junior high school sample and the availability of calculus at the senior high level. Figure 3.6 shows the percentages of junior high schools, by SES and race, that offered at least one section of accelerated mathematics.

Despite what appear to be meaningful differences in the percentages of junior high schools offering accelerated mathematics classes—far fewer of the sampled low-SES and predominantly minority schools offered these classes—the differences among all SES categories for accelerated mathematics offerings were not significant.[14] A difference approaching significance, however, was found between the lowest- and highest-SES schools.[15] Similarly, the differences across percentage white categories were not significant,[16] but again, the difference between high-minority and high-white schools approached significance.[17]

Figure 3.7 shows the numbers of sections of accelerated mathematics in schools of different types that offered these courses.[18] The distribution of the number of sections

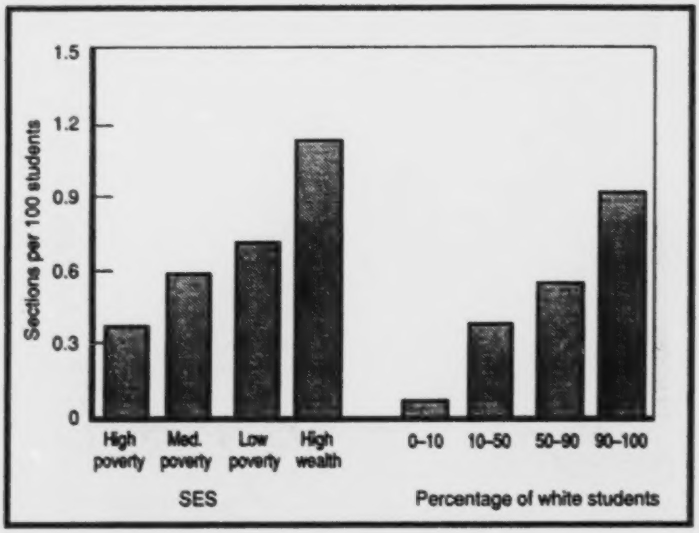


Fig. 3.7—Number of accelerated mathematics classes per 100 students in junior high schools offering accelerated mathematics classes, by school SES and racial composition (for SES, F = 3.06, $P < 0.05$; for racial composition, F = 4.54, $P < 0.01$)

offered is significant across SES categories,[19] and it is highly significant across percentage-white categories.[20] Students attending high-SES or all-white schools have far greater opportunities to take critical gatekeeping courses that will prepare them for advanced mathematics and science courses in senior high school than students in low-SES and high-minority schools.

Senior High School. We were also interested in the percentage of senior high schools offering calculus classes.

Figure 3.8 shows the striking difference in students' opportunities to take calculus at schools of different SES levels. Predominantly white schools were found to be far more likely to offer calculus than high-minority schools,[21] although the differences across all four racial groupings did not reach statistical significance. Eighty percent of the predominantly white schools offered at least one section of calculus, whereas only about 50 percent of the high-minority schools did. This means that students at low-SES

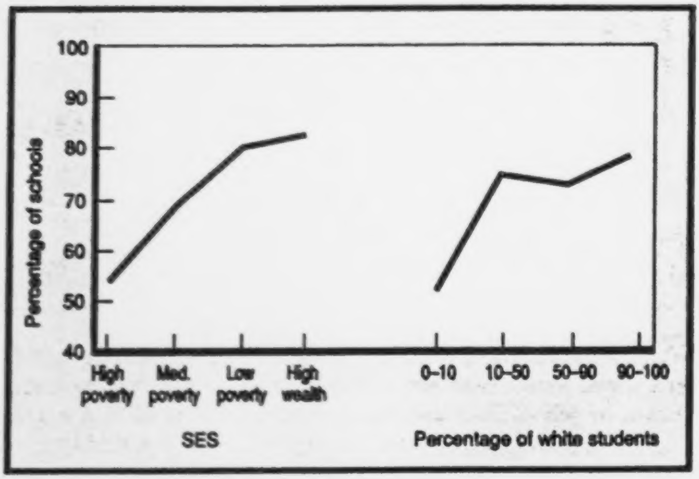


Fig. 3.8—High schools offering calculus classes, by school SES and racial composition (for SES, F = 4.21, P < 0.01; racial composition had no significant effect overall, although groups 1 and 4 differed with P < 0.05)

or high-minority schools who are prepared to take calculus in high school have far less opportunity to do so than their peers at economically more advantaged or predominantly white schools.

Among schools offering calculus, a further dimension of opportunity is reflected in the number of sections offered relative to the size of the student body in various types of schools. Figure 3.9 shows that the opportunity to take calculus increases considerably as the proportion of low-income students drops, and that while mixed schools have relatively comparable numbers of calculus offerings, high-minority schools that offer calculus have far fewer sections. Moreover, the differences in students' access to calculus are minimized in these figures by the omission of schools that offered *no* sections.

It is difficult to determine exactly why schools offer the types of courses they do. Schools that offer few rigorous courses often do so because they have few students

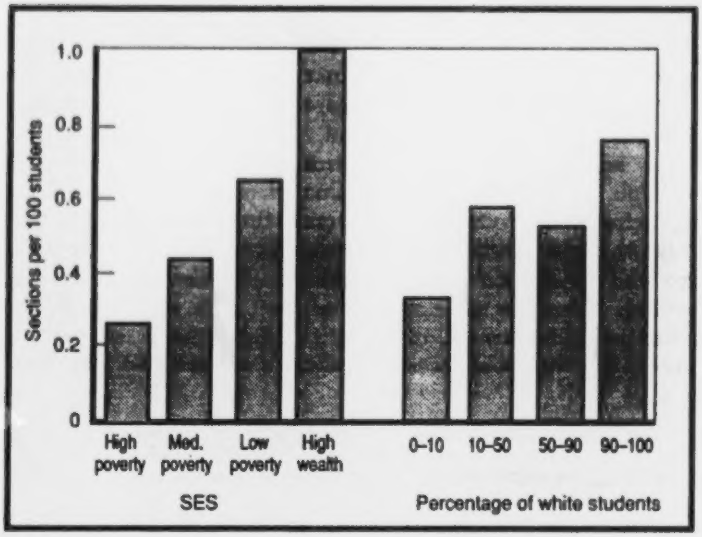


Fig. 3.9—Number of calculus classes per 100 students, by school SES and racial composition (for SES, F = 11.28, $P < 0.001$; for racial composition, F = 3.05, $P < 0.05$)

who are "qualified" to take those courses, that is, who meet test score or other criteria conventionally seen as prerequisites for learning content such as algebra, geometry, advanced mathematics, or mathematics-related science subjects. In fact, the most widely accepted explanation is that secondary schools' programs are constrained by earlier conditions that fail to develop the skills of disadvantaged students. But it should be noted that some schools bend conventional placement criteria and encourage lower-scoring students to take rigorous science and mathematics courses.

Some schools may simply place less emphasis on academic courses or, alternatively, more emphasis on more practical offerings, in the belief that such courses better meet their students' needs. Still other schools may lack staff who are able to teach certain courses, and this may be a significant factor in the paucity of curricular offerings in high-minority and high-poverty schools. Whatever the reasons, secondary schools serving low-income, minority, and inner-city populations offer fewer sections of college-preparatory or advanced courses and more general, applied, and remedial courses. The relatively small differences in the *numbers* of class sections offered in high school science and mathematics programs obscure substantial differences in the depth and rigor of programs at different types of schools.

Moreover, differences in students' opportunities to take critical gatekeeping courses affect their access to continuing study in science and mathematics. Of course, some qualified students, especially in cities, can transfer to schools that offer advanced programs. In these cases, students' opportunities may not be as restricted as our data suggest. However, this does not contradict the clear finding that fewer schools serving low-income and minority students offer the opportunity to begin the college-preparatory mathematics sequence in junior high school or to take calculus in senior high school. And among schools that offer these courses, the proportion of students who take them (as electives or as requirements) is far greater at high-income and predominantly white schools. These findings are quite disturbing: They are not simply consistent with the typically lower mathematics achievement of minority high school students, they strongly signal unequal access to valuable science and mathematics knowledge.

Access to Courses Within Schools

Schools' judgments about students' abilities and their likelihood of going to college affect access to various course offerings and shape the paths students take through the school curriculum. In most schools, fewer mathematics and science courses are available to low-track students or required for them (Guthrie and Leventhal, 1985; Sanders, Stone, and LaFollette, 1987; Vanfossen, Jones, and Spade, 1985).[22]

At the NSSME schools, traditional academic or standard college-preparatory courses and advanced college-preparatory courses are offered most often to students perceived as having high ability, somewhat less often to students thought to be of average ability, and hardly ever to students seen as having low ability (Table 3.2). Low-track students more often take general or applied science and mathematics courses.[23] While some of these courses include basic or applied "versions" of standard academic subjects (e.g., pre-algebra or introduction to algebra, general biology, applied chemistry, basic or "fun" physics), the content in most is quite different from the standard fare in academic science and mathematics.

Table 3.2

DISTRIBUTION OF GENERAL, ACADEMIC, AND ADVANCED SCIENCE AND MATHEMATICS COURSES IN SENIOR HIGH SCHOOLS, BY ABILITY LEVEL OF CLASS

(Percentage of courses of each type taken)[a]

	Type of Course			
Class Ability Level	General	Academic	Advanced	Total
Predominantly low	74	25	1	100
Predominantly average	21	66	12	100
Predominantly high	5	40	55	100

[a]Chi-square = 581.521, $P < 0.001$.

Since low-SES and minority students are disproportionately enrolled in low-track classes and substantially underrepresented in high-ability classes, the likelihood of their being enrolled in academic and advanced courses decreases with the percentage of minority students in the class. Four-fifths of the predominantly white classes in the NSSME sample were taking traditional academic or advanced courses, whereas only 57 percent of the predominantly minority classes were taking such courses. Of course, these differences are shaped in part by the fact that fewer academic courses are offered at high-minority schools.

However, similar differences appear when we control for racial composition of the entire school. As shown in Table 3.3, those classes in which white students are clustered tend to be academic, and those in which minority students are clustered tend not to be.

Just as differences in course offerings at schools of different types result in unequal access to science and mathematics content, curriculum differentiation through ability grouping leads to further race- and class-linked inequalities. Students judged to have low ability (disproportionately large numbers of minority students) have far less access to the standard academic subject matter in science and mathematics than do other students.

Summary

This section has examined how students' access to science and mathematics knowledge varies with their race, social class, and neighborhood, and with the judgments educators make about their intellectual abilities. Except for slightly more time allocated for mathematics instruction in elementary schools with high concentrations of low-income and minority children, access to science and mathematics knowledge is limited for students from groups that consistently achieve and participate less in these areas. To the extent that they are enrolled in secondary schools where they are the majority, low-income students, African-Americans, and Hispanics have less-extensive and less-demanding science and mathematics programs available to them, and they have considerably fewer opportunities to take the critical gatekeeping courses that prepare them to pursue science and mathematics study after high school. This disadvantage is

Table 3.3

DISTRIBUTION OF GENERAL, ACADEMIC, AND ADVANCED SCIENCE AND MATHEMATICS CLASSES, BY CLASS RACIAL COMPOSITION

(Percentage of courses of each type taken)[a]

Class Racial Makeup	Type of Course General	 Academic	 Advanced	Total
Disproportionately minority	59	36	5	100
Racially proportionate (±10 percent)	23	51	25	100
Disproportionately white	15	49	36	100

NOTE: For these analyses, we considered only those schools with white populations greater that 10 percent, but less than 90 percent.

[a] Chi-square = 107.414, $P < 0.001$.

compounded by differences in students' opportunities within schools. Students who are thought to be of low ability are far less likely to be placed in traditional academic courses than are students judged to be more capable. These disparities undoubtedly reflect earlier and broader conditions that fail to develop skills of large numbers of disadvantaged students rather than overt discrimination in the course enrollment process. But the net effect is that disadvantaged and minority students have considerably less access to knowledge that is considered necessary either for science and mathematics careers or for becoming scientifically literate, critically thinking citizens and productive members of an increasingly technological workforce.

Schools in the United States ration curriculum far more than those in many other countries. The Second International Math Study (SIMS) found that U.S. middle schools and junior high schools routinely sort eighth graders into four types of mathematics classes (remedial, typical, enriched, and algebra); this ability grouping was more extensive than that practiced in any other country studied. The SIMS also found that different types of classes provided students with quite different access to mathematics topics. For example, only the small percentage of U.S. students who were enrolled in algebra classes spent much of their class time studying algebra topics. In contrast, Japanese schools exposed nearly all their seventh graders to an intensive algebra curriculum. Only 13 percent of U.S. seventeen-year-olds were enrolled in advanced mathematics, and only about 20 percent of these advanced classes included calculus. This contrasts with many other countries where far greater percentages of students were enrolled in advanced mathematics (e.g., Hungary, Japan, Canada, Ireland, Scotland, Sweden, Thailand), nearly *all* of them in calculus courses (Travers and Westbury, 1989). Most surprising, the more exclusive populations of U.S. students who were taking calculus in the twelfth grade did not outscore the broader groups from other nations on advanced mathematics tests (McKnight et al., 1987). These data suggest that this highly selective system does not enhance achievement, even for those students enrolled in the most advanced courses.

Notes

1. See Carey (1989) for a review of the literature on this topic.

2. Teachers reported the number of days per week they typically presented lessons in science and mathematics and the appropriate number of minutes they spent in an average lesson. To obtain an estimate of the average minutes per day in each subject, we multiplied the reported minutes per lesson by the number of lessons per week and divided the result by 5.

3. $F = 5.05$, $P < 0.05$.

4. While we do not know with certainty that the homogeneous classes in our sample were groups of students who remained together for the whole day or groups of students pulled from heterogeneous homerooms for science and mathematics instruction, we suspect that many were the former. We base our speculation on the fact that many elementary schools regroup students from heterogeneous classes for mathematics instruction, but few do so for science (Oakes, 1985; Slavin, 1987). However, since only slightly smaller percentages of science than mathematics classes were identified as homogeneous (63 percent of science and 70 percent of mathematics), we suspect that most of the homogeneous classes in the sample stayed together for most or all of the school day.

5. Because the survey reported only how much time the classroom teacher spent on lessons, we do not know the extent of additional time low-track students might have spent in pull-out, remedial programs that supplemented regular classroom instruction.

6. Class sections refers to the total number of classes offered by the school, not the number of discrete course offerings. For example, two schools may both offer 3 science courses: biology, chemistry, and physics. However, the first school may offer 3 sections of biology, 2 of chemistry, and 1 of physics, while the other offers 6, 5, and 2, respectively. While both schools offer 3 courses, the first offers 6 sections and the second offers 13. The importance of this is that it suggests greater participation rates for individual students if overall school enrollments are equivalent.

7. The secondary and middle schools surveyed in the NSSME included schools with a variety of grade spans (9-12, 10-12, 7-9, 6-9, etc.). In this study, we usually used subsets of these schools. The 9-12 and 10-12 schools as a group represent high schools, and schools with maximum grades 8 or 9 represent junior high schools. The high school subgroup comprised about 360 schools; the junior high subgroup, roughly 200 schools. In some cases, we had to leave off the sampling weights because the schools were originally assigned a "senior" or a "junior" weight, but not both. This became a problem, since some of the 9-12 schools were classified as junior highs (those in which only ninth grade teachers were sampled). Thus, for some of our school-level analyses, many of the more inclusive senior high schools (e.g., 9-12 schools) would not have the appropriate weights.

8. To account for the size of the school and its grade span, which might affect the number of courses offered per student, we used an analysis of covariance to control for the logarithm of number of students per grade and the minimum and maximum grades in school.

9. $P < 0.05$.

10. The subset of schools we used comprised all the schools with grade spans of 9-12 and 10-12. However, we considered only schools with more than 30 students in a grade to avoid the distortions extremely small schools can create. To account for school size and the presence or absence of a ninth grade (which can affect types of science offerings, for example), our model included the logarithm of the number of students per grade and the minimum grade in a school. After controlling for these variables, we found that the distributive pattern by SES categories was significant at the 0.05 level for total science courses ($F = 3.27$, $P = 0.02$), but not for total mathematics courses ($F = 2.06$, $P = 0.11$). Racial composition (percentage white) did not result in significant differences with respect to total course offerings in either mathematics or science

($F = 1.22$, $P = 0.30$ for science; $F = 0.34$, $P = 0.80$ for math). Significant differences in mathematics programs were found in schools in different types of communities, however, with rural schools having the least-extensive programs ($F = 3.05$, $P < 0.01$).

11. The courses that were classified in each of these categories are listed in the Appendix.

12. $F = 3.03$, $P < 0.05$; and $F = 3.63$, $P < 0.05$, respectively.

13. The subsample of schools examined for calculus offerings was identical to that in the general mathematics and science offerings analysis, but the accelerated-mathematics analysis is based on our "junior high" grouping. The count of accelerated mathematics classes was determined by the number of algebra classes offered (in schools whose final grade is 8), or by the number of geometry classes offered when algebra was also offered (in schools with a ninth grade).

14. $F = 1.30$, $P = 0.27$.

15. $P = 0.09$.

16. $F = 1.04$, $P = 0.38$.

17. $P = 0.08$. The lack of significance of these results may be attributable, in part, to the small number of schools in some of the categories.

18. As in the earlier analyses, course offerings were divided by the average number of students per grade.

19. $F = 3.06$, $P < 0.05$.

20. $F = 4.54$, $P < 0.01$.

21. $P > 0.05$.

22. However, access to courses and coursetaking varies among the schools with different tracking policies (Rosenbaum, 1986; Oakes, 1985). For example, students enrolled in nonacademic tracks in Catholic high schools typically take more academic courses and fewer electives than do their counterparts in public schools; the Catholic schools also place more students in academic tracks (Lee, 1986; Lee and Bryk, 1988).

23. Business mathematics, consumer mathematics, remedial mathematics, general mathematics, technical mathematics, applied mathematics, basic or remedial science, general science, agriculture, current issues, medical technology, plant science, electronics, aviation, ecology, environmental science.

References

Barr, R., and Dreeben, R. (1983). *How Schools Work*. Chicago: University of Chicago Press.

Berliner, D.C. (1979). "Tempus Educare." In *Research on Teaching: Concepts, Findings, and Implications*, edited by P.L. Peterson and H.L. Walberg. Berkeley, CA: McCutcheon Publishing.

Brophy, J.E., and Good, T.L. (1986). "Teacher Behavior and Student Achievement." In *Handbook of Research on Teaching*, edited by M.C. Wittrock, 328-375. New York: MacMillan.

Carey, N. (1989). "Instruction." In *Indicators for Monitoring Mathematics and Science Education: A Sourcebook*, edited by R.J. Shavelson, L.M. McDonnell, and J. Oakes. Santa Monica, CA: The RAND Corporation.

College Entrance Examination Board. (1985). *National College Bound Seniors, 1985*. New York: College Entrance Examination Board.

Crosswhite, F.J., Dossey, J.A., Swafford, J.O., McKnight, C.C., and Cooney, T.J. (1985). *Second International Mathematics Study: Summary Report for the United States*. Washington, DC: American Association for the Advancement of Science.

Darling-Hammond, L. (1985). *Equality and Excellence: The Educational Status of Black Americans*.

New York: College Entrance Examination Board.

Ekstrom, R.B., Goertz, M.E., and Rock, D. (1988). *Education & American Youth*. Philadelphia, PA: The Falmer Press.

Gamoran, A. (1987). "The Stratification of High School Learning Opportunities." *Sociology of Education*, 60, 135-155.

Goodlad, J.I. (1984). *A Place Called School: Prospects for the Future*. New York: McGraw Hill.

Grant, C.A., and Sleeter, C.E. (1986). "Race, Class, and Gender in Education Research: An Argument for Integrative Analysis." *Review of Educational Research*, 56(2), 195-211.

Guthrie, L.F., and Leventhal, C. (1985). "Opportunities for Scientific Literacy for High School Students." Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Hallinan, M.T., and Sorensen, A.B. (1983). "The Formation and Stability of Ability Groups." *American Sociological Review*, 48, 838-851.

Husen, T. (Ed.) (1967). *International Study of Achievement in Mathematics: A Comparison of Twelve Countries* (Vols. 1 and 2). New York: John Wiley.

Jones, L.V. (1984). "White-Black Achievement Differences: The Narrowing Gap." *American Psychologist*, 39, 1207-1213.

Jones, L.V., Burton, N.W., and Davenport, E.C. (1984). "Monitoring the Achievement of Black Students." *Journal for Research in Mathematics Education*, 15, 154-164.

Jones, L.V., Davenport, E.C., Bryson, A., Bekhuis, T., and Zwick, R. (1986). "Mathematics and Science Test Scores as Related to Courses Taken in High School and Other Factors." *Journal of Educational Measurement*, 23(3), 197-208.

Lee, V.E. (1986). "The Effect of Curriculum Tracking on the Social Distribution of Achievement in Catholic and Public Secondary Schools." Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Lee, V.E., and Bryk, A.S. (1988). "Curriculum Tracking as Mediating the Social Distribution of High School Achievement." *Sociology of Education*, 62, 78-94.

McKnight, C., Crosswhite, F.J., Dossey, J.A., Kifer, E., Swafford, J.O., Travers, K.J., and Cooney, T.J. (1987). *The Underachieving Curriculum*. Champaign, IL: Stipes Publishing.

National Science Board, Commission on PreCollege Education in Mathematics, Science, and Technology (1983). *Educating Americans for the 21st Century: Source Materials*. Washington, DC: National Science Foundation.

Oakes, J. (1987). "Tracking in Secondary Schools: A Contextual Perspective." *Educational Psychologist*, 22, 129-154.

Oakes, J. (1985). *Keeping Track: How Schools Structure Inequality*. New Haven, CT: Yale University Press.

Oakes, J. (1983). "Limiting Opportunity: Student Race and Curricular Differences in Secondary Vocational Education." *American Journal of Education*, 91, 801-820.

Oakes, J., Gamoran, A., and Page, R. (in press). "Curriculum Differentiation." In *Handbook for Research on Curriculum*, edited by P.W. Jackson. New York: MacMillan.

Orland, M.E. (1988). "The Demographics of Disadvantage: Intensity of Childhood Poverty and Its Relationship to Educational Achievement." In *Access to Knowledge*, edited by J. Goodlad and P. Keating. New York: College Entrance Examination Board.

Pallas, A.M., and Alexander, K.L. (1983). "Sex Differences in Quantitative SAT Performance: New Evidence on the Differential Coursework Hypothesis." *American Educational Research Journal*, 20, 165-182.

Peng, S.S., Owings, J.A., and Fetters, W.B. (1981). "Effective High Schools: What Are Their Attributes?" Paper presented at the 1981 Annual Meeting of the American Statistical Association, Cincinnati, OH.

Persell, C.H. (1977). *Education and Inequality: The Roots and Results of Stratification in America's Schools*. New York: The Free Press.

Rock, D., Braun, H.I., and Rosenbaum, P.R. (1985). *Excellence in High School Education: Cross-sectional Study, 1980-1982. Final Report*. Princeton, NJ: Educational Testing Service.

Rock, D., Braun, H.I., and Rosenbaum, P.R. (1984). *Excellence in High School Education: Cross-sectional Study, 1972-1980. Final Report*. Princeton, NJ: Educational Testing Service.

Rosenbaum, J.E. (1986). "Institutional Career Structures and the Social Construction of Ability." In *Handbook of Theory and Research for Sociology of Education*, edited by J.G. Richardson. New York: Greenwood Press.

Rosenbaum, J.E. (1980). "Social Implications of Educational Grouping." *Review of Research in Education*, 8, 361-401.

Rosenholtz, S., and Simpson, C. (1984). "The Formation of Ability Conceptions: Developmental Trend of Social Construction." *Review of Educational Research*, 54, 31-63.

Sanders, N., Stone, N., and LaFollete, J. (1987). *The California Curriculum Study: Paths Through High School*. Sacramento, CA: California State Department of Education.

Sells, L. (1982). "Leverage for Equal Opportunity Through Mastery of Mathematics." In *Women and Minorities in Science: Strategies for Increasing Participation*, edited by S.M. Humphreys. Washington, DC: American Association for the Advancement of Science.

Slavin, R.E. (1987). *A Review of Research on Elementary Ability Grouping*. Baltimore, MD: Johns Hopkins University Press.

Slavin, R.E., Braddock, J.H., Hall, C., and Petza, R.J. (1989). *Alternatives to Ability Grouping*. Baltimore, MD: Center for Research on Effective Schooling for Disadvantaged Students.

Travers, K.J., and Westbury, I. (1989). *The IEA Study of Mathematics I: Analysis of Mathematics Curriculum*. Oxford: Pergamon.

Vanfossen, B.E., Jones, J.D., and Spade, J.Z. (1985). "Curriculum Tracking: Causes and Consequences." Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Weiss, I. (1987). *Report of the 1985-86 National Survey of Science and Mathematics Education*. Research Triangle, NC: Research Triangle Institute.

Welch, W.W., Anderson, R.E., and Harris, L.J. (1982). "The Effects of Schooling on Mathematics Achievement.." *American Educational Research Journal*, 19, 145-153.

Wolf, R.M. (1977). *Achievement in America*. New York: Teachers College Press.

WHY GIRLS DON'T KNOW

Jane Butler Kahle

In a 1987 study, researchers approached the results of the 1976-77 National Assessment of Educational Progress (NAEP) for science from a different angle. Rather than looking at only correct and incorrect answers, M.C. Linn, T. de Benedictis, K. Delucchi, A. Harris, and E. Stage also examined the "I don't know" responses. When they compared boys and girls, they found that much of the achievement differences could be accounted for by differences in answering patterns. Girls, significantly more often than boys, selected the "I don't know" response.

Why do so many girls respond, "I don't know?" Are they less knowledgeable, less informed? Or are they more honest, more modest? If girls are to participate more in the sciences, these questions need to be answered. Recent studies of classroom practices and teacher intervention programs suggest some answers to these problems. In order to understand what teachers can do, I will assess the situation and propose some solutions. I will begin with a review of today's science classrooms, elementary through college, followed by a survey of recent programs and their solutions.

Assessing the Situation

According to E.F. Keller (1986), gender is "what a culture makes of sex. It is the cultural transformation of male and female infants into adult men and women" (p. 122). My analysis of *Why Girls Don't Know* is an interpretation of gender differences; that is, differences in interactions not attributable to biological differences, which begin before school and continue throughout life. For example, even though preschool girls and boys are physically similar, they already have very different experiences. Boys' experiences provide them with backgrounds, interests, and attitudes which are important for later achievement in science (Kahle and Lakes, 1983; Sjoberg, 1986). They handle more tools, throw more balls, construct more Lego bridges, build more block towers, and tinker more with simple mechanical objects. In addition, boys generally receive more subtle rewards for taking risks. Exclamations such as, "What a brave boy!" or "Isn't he strong!" reward a boy for climbing trees or jumping into pools. A girl, on the other hand, is praised for "being a little lady" and "keeping her pretty dress clean." While these differences are subtle, they provide one sex with an appropri-

---

*This article first appeared in **What Research Says to the Science Teacher—The Process of Knowing** (1990), Volume 6, pp. 55-67. Washington, DC: National Science Teachers Association.*

ate background, as well as the necessary attitudes, for success in science. As S. Johnston (1984) observes:

> Boys and girls enter school science classrooms with different past experiences, different interests, different attitudes and different expectations. This indicates that teachers cannot dismiss the problem of girls' under-achieving in science by treating boys and girls identically...[T]he science classroom and curriculum are designed to build on a foundation of interests, experience and attitudes that is present for one sex but not for the other. Treating boys and girls identically in school can serve to accentuate rather than diminish the existing differences (p. 22).

Unfortunately, evidence suggests that girls do not even receive equal treatment in schools, not to mention the kind of treatment needed to encourage them in the sciences. Nothing in the curriculum acts to counter girls' lack of out-of-school science experiences, and this inexperience continues to be a problem throughout elementary and secondary school and into college.

Elementary School

In assessing elementary classrooms, studies indicate that boys and girls bring different science experiences to school, and once in school they receive very different science educations. Studies in England (Smail, 1984), the United States (Kahle and Lakes, 1983), Norway (Jorde and Lea, 1987), and Australia (Parker and Rennie, 1986) document very different behaviors by boys and girls in the elementary classroom. Fewer girls than boys handle science equipment, perform science experiments, or participate in science-related activities. The different backgrounds that boys and girls bring to elementary school are perpetuated by the schools; for example, I. Mullins and L. Jenkins (1988) report that only 37 percent of girls in the third grade, compared with 51 percent of the boys, have used microscopes. Equal numbers of girls and boys may sit through science lessons, but they participate in them in unequal ways.

As W. Harlen (1985) states:

> The opportunity for children to experience science activities exists at the elementary-school level perhaps more readily than at later stages of education. If we wish to increase girls' access to science therefore, science at this early level has a vital part to play (p. 545).

The elementary school is the critical place for change: change in formal and informal science curricula, change in classroom instruction and interactions, and change in school structure and socialization.

When children first enter school, international projects report that although boys and girls express interest in slightly different types of science, their overall levels of interest are similar.

Furthermore, for both boys and girls, interests are directly related to areas of daily experience. For example, L.H. Parker and L.J. Rennie (1986) report the following:

> Where the experience is one which is likely to be universal (e.g., earthworms, shadows, germs, and water), very little sex-differentiation of interest is shown. In other areas, such as wheels and motors and growing a vegetable or flower garden, where boys' and girls' out-of-school experiences are likely to be quite different, clear sex-stereotyping is revealed (p. 177).

Sex-related differences in out-of-school experiences may be reinforced within schools if children are allowed to select science topics based on interest. Therefore, the elementary science curriculum is extremely important. It must be motivating, yet include a range of topics; it must include opportunities to handle the tools of science, yet develop a conceptual foundation for later studies.

The student-centered, activity-based curricula of the 1960s and 1970s fulfill the criteria for an ideal curriculum. Ironically, they are little used, and when they are implemented, they are frequently misused. These curricula, SCIS (Science Curriculum Improvement Study), SAPA (Science, A Process Approach), and ESS (Elementary Science Study) have faded from classrooms for three primary reasons: Teachers have not understood the scientific principles the materials promulgated, classrooms have not been organized for small group interaction in science, and schools have not provided the equipment or the scheduling required. However, recent analyses make a strong case for experience-based curricula. For example, a review of 34 evaluative studies indicated that children using the "hands-on" curricula performed better on every measure of achievement than children studying "textbook" science. In fact, one analysis of 13,000 children in 1,000 U.S. classrooms has demonstrated that children instructed using activity-based materials surpass those who received traditional instruction. They score higher on tests of science processes, creativity, perception, logic, development, science content, and mathematics. In addition, the experienced-based curricula provide opportunities for experimentation as well as for handling instruments, making measurements, observing natural phenomena, collecting data, and making interpretations. These curricula also have the potential of producing equal outcomes for both sexes. As C. Iliams (1985) states, "girls are less likely than boys to make up the education deficiency in out-of-school experiences" (p. 79), therefore a "hands-on" curriculum is necessary in order to provide opportunities for the experimental background which girls lack when they enter school.

Although activity-based curricula may provide needed experiences, teachers must implement them carefully to insure equal opportunity for participation and equal expectation of performance by both sexes. Teachers may need to reorganize classroom activities to provide girls with the extra time and opportunities to do science in order to get the same level of performance that teachers anticipate and encourage from boys. Another obstacle to equal education opportunity is that girls frequently choose not to participate in science demonstrations and experiments in spite of encouragement from teachers. Because socialization discourages girls from science, teachers fear that forcing involvement may result in increasingly unfavorable attitudes. In a 1976-77 study of science attitudes, however, NAEP revealed that girls' participation levels may not reflect their attitudes. NAEP asked 9-year olds a series of questions which were phrased in two ways: Children were asked, "Would you like to . . ." (desired participation), and "Have you ever . . ." (actual participation) for the same activities (Kahle and Lakes, 1983). Girls' actual levels of participation were lower than their expressed desire to participate in many activities.

Girls often do not receive the encouragement that boys do. Instead, elementary teachers tend to react more favorably to boys. When these teachers are asked to identify scientifically talented or gifted students, a cross-cultural pattern emerges: They identify more boys. Furthermore, observation of both the number and duration of teachers' interactions with identified creative girls and boys revealed that teachers interact twice as often and for longer durations with boys, compared to girls. In England, M.G. Spear (1984) analyzed the markings of science papers attributed to 12-year-old boys and girls and found that both male and female science teachers generally gave higher marks when the work was attributed to a boy. In addition, different interaction patterns are found. Frequently, teachers allow boys to dominate discussions and the use of equipment, and they are four times more likely to select boys as target students than girls (Tobin and Garnett, 1987).

The elementary school as an institution also provides different educational experiences for girls and boys. Current institutional practices consistently use gender as the basis to group children when many other criteria are available. As an alternative, school activities could be designated as quiet or adventurous with all students taking part in the different kinds of activities in the playground and the classroom. By ensuring equity in the corridors as well as in the classroom, schools would come much closer to providing optimal settings for science instruction.

Secondary School

Although equal numbers of boys and girls attend science classes in elementary and middle/junior high schools where enrollment is mandatory, in high school the situation changes. After age 14, the differences in out-of-school science experiences for boys and girls are exacerbated by different in-school science experiences (Kahle and Lakes, 1983). Although equal numbers of boys and girls take high-school biology, only 30 percent of high-school girls, compared with 39 percent of boys, take chemistry. Physics, taken by 26 percent of all high-school boys, is studied by only 14 percent of the girls.

Furthermore, the ratio of male-to-female science students is reflected in the ratio of male-to-female science teachers. In the United States, only 24 percent of secondary-school science teachers are women, and the majority of them teach biology (Weiss, 1978). Without role models, girls may feel less comfortable or more out-of-place in upper-level science courses.

The decline in the number of girls continuing on in physical science courses (chemistry and physics) has been attributed to many causes: lack of interest, lack of aptitude, and science and/or math anxiety. In addition, both curricula and teaching styles may affect enrollment patterns of boys and girls in elective science courses.

In hopes of increasing the appeal of science to women and minorities, the publishing industry has increased pictorial representations of female and minority populations. Publishing guidelines ensure that segments of our population are represented pictorially in correct proportions, and therefore, books now have approximately 50 percent of illustrations and diagrams showing females and 17 percent depicting Blacks. While these cosmetic changes are a beginning, textbooks still need substantive changes. For example, in the 1985 editions of two popular high-school biology texts, between 75 percent and 98 percent of the cited scientific work described the contributions of men, while the texts cited the contributions of women as only 2 percent and 4 percent of the total. In spite of the changes thus far, women's role in science still seems minimal and uninviting for young girls today.

Teaching styles also may affect enrollment patterns. In 1981, M. Galton identified three different styles in science teachers: problem solvers, informers, and enquirers. The *problem solver* teacher uses many questions and a low frequency of pupil-initiated and maintained activities; the *informers* primarily use teacher delivery of facts and infrequently use questions except to recall facts; and *enquirers*, who use pupil-initiated and maintained experiments as well as inferring, formulating, and testing hypotheses. Generally, girls prefer the latter style of teaching, the enquirers. It comes as no surprise that this is the style most often used in biology classes, which girls often select. In contrast, the problem-solver strategy is least favored by girls, and it is more frequently used in physics, which few girls elect to study. In addition, female teachers more commonly use the enquiry style, while men favor the informer or problem-solver mode of instruction (Douglass, 1985). Could the different teaching styles influence girls away from physics and chemistry?

Classroom ethnographers also divide science instruction into several other categories: whole-class (lecture/demonstration), whole-class interactive (discussion), small group (laboratory), and individual (seat work/projects). Girls and boys respond differently to these various modes of instruction, yet overwhelmingly teachers use the whole-class mode. For example, L. Trowbridge and R. Bybee (1987) report that high-school teachers lecture over 70 percent of instructional time, and K. Tobin's (1987) research in the United States and in Australia indicates that the vast majority of secondary teachers studied (29 out of 33 in his studies) primarily use whole-class instruction. Tobin reports certain differences in teachers' interactive patterns with boys, as opposed to girls, during whole-class instruction.

> Teachers tended to involve males and females to an equal extent in lower cognitive level interactions, but tended to involve males to a greater extent than females in higher cognitive level interactions. Males also participated in a more overt way than females by volunteering to respond to teacher questions by raising their hands when teachers asked questions in a whole-class setting. The major consequence of this engagement pattern was that "target" males were involved in responding to questions intended to stimulate thinking or to elicit responses that would provide a bridge to a new area of content . . .The pattern of male students being more involved than females in whole-class interactions was apparent in classes taught by male and female teachers (p. 39).

Tobin and others have found similar patterns in classes that are tracked on science ability and in ones that are elected by advanced students. For example, in one eleventh-grade biology class which contained 12 girls and 9 boys, approximately 70 percent of teacher questions were answered by boys. In all cases, the observed teachers invariably claimed that they involved girls and boys equally.

Tobin (1987) reports few gender differences in teacher-student interaction patterns during individualized activities; however, gender differences still occur during laboratory activities and in the classroom. Tobin, concurring with J. Whyte (1986), reports that boys tend to dominate the equipment as well as the activities. Whyte concludes that "boys are pushier and seem to regard scientific or technical resources as rightfully theirs. Girls, for their part, become unwilling to enter into an undignified scrabble for equipment" (p. 34). As noted, boys are also asked more higher-order cognitive questions than are girls (Tobin and Garnett, 1987; Tobin and Gallagher, 1987). In the classroom, teachers act differently toward boys and girls. Boys are urged to "try harder" when they do not succeed (Sadker and Sadker, 1985). When giving instructions for completing a problem, teachers give boys specific instructions, while they may show girls how to finish a task or do it for them. They subtly convey to students that boys have the ability to succeed in science and mathematics, but girls do not. A recent study suggests that this non-verbal message may account for a large portion of the gender differences observed in science achievement, since it causes girls to lose self-confidence in their scientific ability. Does the "I don't know" response, discussed at the article's beginning, reflect this lack of self-confidence? Linn and her coworkers found that in a study involving 13- and 17-year olds, girls tend to use the "I don't know" response, particularly for physical-science items or for items with masculine references. They reported that gender-related differences in responses are due to lack of confidence and to differences in prior instruction (Linn et al., 1987). Teacher behavior, therefore, has an important effect on confidence levels as well as on achievement levels.

As noted previously, girls, compared with boys, have fewer verbal exchanges with teachers. This pattern continues even after teachers are sensitized to the difference. For example, S. Hyde (1986) reports that although one of her primary goals was to interest girls in physics, when her student-teacher interaction patterns in her classes were recorded, they show that she still spent 82 percent of her time with boys. What happened? One subtle possibility is that even though some teachers learn to monitor their verbal behavior patterns, they may continue to nonverbally communicate different expectations. For example, studies document that teachers use the following strategies to indicate anticipation of a superior performance: leaning forward, looking into eyes, nodding, and smiling. Do girls receive these non-verbal cues of anticipation?

The critical problem with different expectations is that they are group based, and therefore achievement expectations become a function of one's sex. Because students cannot change their gender, they accept the achievement expectation as something they cannot change. Therefore, different teacher expectations for boys and girls in certain subjects (for example, science and reading) contribute to the cultural expectations from different biological sexes and lead to a gender difference. For example, J.A. Rowell (1971) reports that teachers who expected girls to have difficulty learning physics had girls who did not perform as well as boys—a well-documented gender difference. However, no achievement differences were found between girls and boys who were enrolled in physics classes taught by teachers who did not hold such views. Clearly, teachers can either contribute to or detract from the achievement of girls in science.

Colleges and Universities

S.E. Berryman (1983) has shown that the attrition of women from science increases near the end of their education; that is, during the undergraduate and graduate years. Only 14 percent of the women entering college intend to major in science or engineering, compared to 40 percent of male freshmen (Vetter, 1987).

Despite increases in the numbers of women majoring in science and engineering during the 1970s, some areas peaked in 1984 and have begun to decline. For example, the percentage of women receiving a bachelor's degree in math, computer science, or statistics fell from 4.4 percent in 1986 to 2.4 percent in 1989 (in actual numbers from 20,400 to 9,600) (Vetter, 1987). Furthermore, a higher proportion of both undergraduate and graduate women, compared with men who elect a scientific course of study, chose to leave it before receiving a degree (Matyas, 1986; Hite, 1983).

Recently, several researchers have assessed the factors which lead to higher attrition rates for women, compared with men, from college science majors. N.C. Ware and V. Lee (1987) have summarized the research and identified six factors:

- Girls complete fewer high-school courses in science and math than boys.
- Girls demonstrate less quantitative ability as evidenced by the math section of the Scholastic Aptitude Test.
- Girls have lower self-estimates of their mathematic and scientific abilities.
- Girls perceive discrepancies between personality characteristics associated with femininity and those associated with scientists.
- Girls' parents have lower aspirations for them than boys' parents.
- Women in college tend to react more negatively to initial college courses in mathematics and science than comparable men.

Only one of the factors identified by Ware and Lee relates to academic ability or to a gender-related cognitive difference; that is, the lower average scores of girls, compared to those of boys, on the math section of the SAT. More recent studies, however, indicate that the average difference in girls' and boys' quantitative SAT scores is directly related to the number of math and math-related courses taken in high school rather than innate ability (LeBold, 1987). Ware and Lee analyzed the *High School and Beyond* data for college science majors. They found that high-school counselors and teachers positively influenced boys to take math and science but negatively influenced girls to do so, thus contributing to girls' lower math scores on the SAT.

Ware and Lee also found that personal and social factors, rather than academic ones, often influenced women against college science majors. They report that "concern for future family and personal life seem to inhibit the choice of a science major for females, but facilitate a science major for males" (p. 20). In other words, only social factors have been identified as contributing to the high attrition rate of women from college science and math courses. No academic or cognitive factors have been found.

Current studies focus on the personal factors which influence many talented women to drop college science majors. In A.L. Gardner's (1986) study, she asked female students in engineering, biology, and nursing to complete a Personal Attributes Questionnaire on which they rated themselves according to masculine and feminine characteristics. She found that only 18.5 percent of the women engineers and 23 percent of the women biologists, compared with 42 percent of the women nursing students, rated themselves as typically feminine. The majority of the female engineering and biology students (62 percent and 64 percent), respectively, selected characteristics which classified them as either masculine or androgynous. Those classifications required high self-ratings on characteristics such as self-confidence and tenacity. However, P. Newton (1986) found that young women enrolled in engineering courses in Britain stressed the feminine aspects of their personalities, perhaps in order to appear to be less different or unusual than women in other courses. The effect of women's self-perceptions on enrollment in and commitment to science and engineering majors needs to be assessed further.

Another factor influencing the high attrition rate of women from science at the college level might be inequities in the classroom. Researchers, skilled in assessing classroom climates, have noted small differential behaviors (micro-inequities) that often occur in the course of everyday exchanges in which individuals are singled out or ignored because of sex. In the classroom, teachers react differently toward their female students than toward their male students. Each incident is trivial, but the cumulative effect maintains unequal treatment based upon a personal characteristic which cannot be changed. Consider the following example:

> "Michelle," Professor Jones announces to his mechanical engineering class, "received the highest mark on the test. Her work is truly impressive and should be a challenge to the rest of you—especially you men."

The unrecognized, but voiced, micro-inequities include:

- the uniqueness of Michelle's achievement
- the challenge to increase competitiveness
- the question if "real men" will allow a woman to win

However, Michelle, herself, is probably pleased with the praise she has received. Only later will she become uncomfortable with some of its probable outcomes. For example, she may be kidded about her grades by the men and that kidding may question her femininity, or she may be excluded from a cooperative learning situation so that she doesn't "set the curve" the next time. The effect of micro-inequities on the choice of and perseverance in science by college women is a continuing and, often, unrecognized problem.

Proposing Solutions

Rather than propose untested solutions, I shall describe the results and promises of current research. The studies span the ages from 9-year olds to 20-year olds, covering science classes from grade school through college. The problem addressed by each is the recruitment and retention of female students in science courses and careers. Beyond recruitment and retention, however, each study focuses on equitable science education in which girls as well as boys participate at high cognitive levels in science classes.

J.B. Kahle (1985) identified and observed biology teachers in the United States who were successful in motivating tenth-grade girls to elect optional physics and chemistry courses. Selected teachers from Maine to California were observed, parents and principals were interviewed, and past and present students were surveyed. Approximately 395 children from varied backgrounds participated in the study, which provided a composite picture as well as a collective pool of data from which commonalities could be identified and generalizations made. It was found that the identified teachers taught in a more individualized way, using many discussions, as well as project and laboratory work. In addition, compared with a national sample control group (Weiss, 1978), they also used more diverse media, field trips, and library research.

In these teachers, researchers found three common practices which were easily applicable to all science teaching. First, teachers who were successful in encouraging students to continue in science quizzed or tested their students once a week. Second, they encouraged creativity, noted by 58 percent of boys and 67 percent of girls; and, third, they fostered basic skill development, according to over 70 percent of both boys and girls surveyed. The researchers also noted that these teachers emphasized skills using quantitative methods, transforming data, interpreting graphs, and visualizing three-dimensional objects (Kahle, 1985). These skills add both to the cognitive development and problem-solving skills of adolescents. The implementation of a science curriculum which develops creativity and originality while fostering cognitive growth would encourage and enable more girls, as well as more boys, to study science.

Studies which have analyzed teacher behaviors suggest that both male and female science teachers need to practice what S.M. Malcolm (1983) calls "directed intervention," in which all students are actively and positively encouraged to participate, respond, and question. If, as A. Kelly (1985) suggests, gender differentiation is maintained by the behavior of the children themselves, rather than teacher behavior, directed intervention may be one of the best ways to improve the cognitive development of girls.

Kelly, Whyte, and B. Smail had a similar goal in their 1984 study in England. The Girls in Science and Technology (GIST) project involved ten comprehensive schools in the Manchester area and studied 2,000 children from the time they entered lower school (age 11) until they made their subject choices at age 14. Based on observations and results, the research team suggested a prototype curriculum and hypothesized about ideal situations in middle/junior high school.

In its search for an ideal curriculum, the GIST project developed and tested new curricula as part of its four-year study; that is, when the teacher-participants requested new or different materials, they were developed. Designed to appeal to all students, particularly to girls, the curricular materials have the following characteristics:

- focuses on relationships as well as rules
- focuses on people as well as machines
- develops a pragmatic rather than a dogmatic approach
- views the world as a network rather than a hierarchy of relationships
- emphasizes the aesthetic as well as the analytical aspects of science
- focuses on nurturing living beings as well as on controlling inanimate things (Smail, 1984, p. 27).

The GIST project proposes the integration of the contributions of female scientists into the curriculum. Many associations such as the National Science Teachers Association can provide references and resources identifying the contributions of female scientists. In addition, GIST stresses the inclusion of "tinkering" activities in school science in order to overcome the lack of such experiences by girls in everyday life. The results of both the GIST project and Kahle's (1985) study demonstrate that the ideal secondary science curriculum must provide experiences rotating three-dimensional figures in space, drawing and conceptualizing three-dimensional forms, and projecting curvilinear distances and outcomes. Such experiences increase a child's visual-spatial ability. Since girls usually have less experience with the toys, games, and activities which enhance spatial ability, opportunities must be constructed in the curriculum. The GIST project reveals that although boys initially score better on the spatial-ability tests, enrolling girls in one technical-craft course eradicates the gender difference (Kelly et al., 1984).

Teachers and researchers concur that extensive laboratory work must be included in the curriculum to equalize the science educations of girls and boys. Laboratory activities can also be used to improve interest and experience in unfamiliar subjects. Perhaps a 15-year-old girl in Louisiana described the interest aspect best when she said, "I enjoy working with microscopes. We had a cow heart and we opened it up. [We] looked in the microscope at the different parts of the inside of the heart and I enjoyed that" (Kahle, 1985, p. 54).

The need for experience with the actual tools and techniques of science is well documented. Boys and girls both express anxiety if they do not have sufficient past experience against which to gauge success. For example, Kahle found that girls express little anxiety about focusing a microscope with which they have had experience but great anxiety about wiring an electric circuit with which they have had none. Similarly, boys express concern about taking the temperature of a living organism, a technique with which they are generally unfamiliar. In order to encourage girls to participate equally in science, an ideal curriculum would present alternative and supplementary materials familiar to girls. By providing examples and models drawn from the common experiences of girls (sewing machines and volleyball) as well as those of boys (cars and football), strides can be made toward equality.

In many countries, definitive studies have shown that experience makes the difference; that is, boys and girls express similar interest in topics with which they both have had experience (Kahle, 1985; Smail, 1985; Parker, 1985). Therefore, both curricula and instructional modes must allow for experiences with the equipment and instruments of science. Such experience may improve the problem-solving abilities of all students and contribute to cognitive development and understanding.

Teacher behavior can also affect student attitude. In a classic experiment conducted by Rennie, Parker, and P.E. Hutchinson (1985) in Western Australia, their results support the efficacy of changing teaching behaviors to effect changes in student attitudes. The researchers provided one set of elementary teachers with intensive training in the skills of teaching electricity as well as with information about non-sexist teaching, while a comparable group received only the skills training. They found "a slight but consistent tendency for students in the Experimental Group classes to perceive girls as more competent with electricity than did students in Control Group classes" (Parker, 1985, p. 12). When they asked the year-five children if they could become electricians, 90 percent of the boys in both the experimental and control classes responded positively, while 85 percent of the girls in the experimental group said "yes" compared to only 70 percent of the girls in the control group (Parker, 1985). The behavioral training the teachers received made a tremendous difference in the attitudes of both the class as a whole and the girls, themselves.

One way of making classroom climates more favorable for girls' cognitive development has been proposed by a Norwegian project. The researcher/teacher observation team for fourth- through sixth-grade science lessons suggests the following change in methodology, which has encouraged more participation from girls in the question/answer sessions with teachers. In the usual classroom presentation, teachers begin a lesson by asking pupils what they know about a topic, and boys answer the majority of the questions based on knowledge from experiences outside the classroom which generally their female counterparts have never had. Instead, if pupils begin the activity at once, with little introduction from the teacher, everyone has access to the experience. A discussion by the teacher after the activity, rather than before, encourages more participation by girls (Jorde and Lea, 1987).

Kahle's (1985) study, discussed earlier, also proposes another means for creating better classroom climates for girls by providing a pleasant, attractive, and stimulating environment for learning science. While the study's pleasant surroundings and activity-initiated lessons provided a relaxed and less competitive atmosphere for studying science, girls' performance levels remained unaffected. Kahle had hoped to increase the girls' interest and performance levels by providing information, female role models, and interest-oriented curricular materials. The project, however, produced limited results. When assessments of students' science achievement, interest, and attitudes were made after the project, it was found that boys, as compared to girls

- had more experiences in science
- indicated higher interest in science careers
- held more positive perceptions of science and scientists
- expressed more positive attitudes toward science and scientists
- found science more useful
- received significantly more A and B grades in science (Kahle, 1987).

The indirect, interest-oriented approach may have improved the biology classroom climate for girls, but it failed to increase the cognitive-achievement levels or to improve the science attitudes of girls. Activities or curricula focused on interests and careers may temporarily encourage girls as well as boys to participate in science, but they will not expand girls' horizons or improve their skills in science. In order to change classroom climates as well as student attitudes and achievement levels, intervention activities must be direct, explicit, and capable of sustaining girls' interest in science. But prerequisite skills must be provided, or girls will lag behind boys in achievement levels and attitudes. Rigorous science experiences and skill development are an essential part of an effective classroom climate.

Rennie et al.'s (1985) study, conducted in elementary schools in Western Australia, combined rigorous skill training with activities designed to improve the classroom climates. They found that more time was spent in group work, and more girls actively participated in the experimental activities after the training sessions. The science classes, therefore, had become more equitable, and the climate had improved for girls.

At the college level, changes in the campus climate are likewise necessary to improve the entrance and retention of women in science. Micro-inequities, as discussed earlier, are usually regarded as a classroom problem, but the atmosphere of the campus may also affect women's selection of non-traditional majors. In a 1986 study, M.L. Matyas assessed the overall campus climates which may affect the attrition of women from college science majors. She compared women and men who were switching from a biology major to other majors and found that the women had higher grade-point averages than the men. The women were capable of the science and mathematics required, but some other factor had motivated the change in major. Matyas found that personal factors often affected the decisions. While some of the personal factors related to future concerns, others related to both campus and classroom climates; for example, a lack of role models, a lack of female friends in a science major, and a male orientation in science.

Many college women in non-traditional majors, therefore, may feel isolated. In W.K. LeBold's (1987) study, he suggests that a "critical mass" of any group is necessary before members of that group feel comfortable in a given situation. "However, when a critical mass is reached, many of the socialization problems are minimized, women recognize that there are others 'in the same boat,' and the resulting self-support groups are eventually institutionalized" (p. 86). LeBold suggests programs that are successful in increasing the number of women in non-traditional fields must address at least four issues: recruitment, retention, future employment, and program evaluation. Furthermore, he warns that fragmented programs, focusing on a single issue, often are ineffective. In addition, LeBold states that single-issue programs may delete or delay any concerted effort to improve classroom or campus climates for women. Rather, a single program may be used to justify the lack of the coordinated, sustained effort needed.

Summary

For over a decade research has focused on the scarceness of women in science. The research has analyzed personal, societal, and educational factors. Data have been collected and statistically analyzed, and yet, the situation has not improved. Girls, in comparison to boys, still do not perform as well on cognitive items in the NAEP's science survey, on the mathematics section of the SAT, and on the Second International Educational Assessments in Science.

Yet, with intervention programs based on current research, teachers can favorably affect the attitudes and achievement levels of diverse students in science. Spatial-visual activities, basic skill development, and tinkering with and testing equipment, as well as frequent tests will help all children progress in science. As more girls succeed and continue in science, a critical mass will be prepared for college science majors. Eventually, more girls will eagerly respond, "Yes, I know" to questions on national and international assessments of science knowledge and skills.

References

Berryman, S.E. (1983, November). *Who will do science?* Washington, DC: The Rockefeller Foundation.

Douglass, C.B. (1985). Discrepancies between men and women in science: Results of a national survey of science educators. In J.B. Kahle (Ed.), *Women in science*. Philadelphia: The Falmer Press.

Galton, M. (1981). Differential treatment of boy and girl pupils during science lessons. In A. Kelly (Ed.), *The missing half*. Manchester: Manchester University Press.

Gardner, A.L. (1986). *Effectiveness of strategies to encourage participation and retention of precollege and college women in science*. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Harlen, W. (1985). Girls and primary-school science and higher education: Sexism, stereotypes and remedies. *Prospects, 15*(4), 541-551.

Hite, L.M. (1983). *A study of doctoral students compared by gender and type of field of study on factors of role congruence, perceived support from faculty, and perceived support from peers*. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Hyde, S. (1986). Girls and science—Strategies for change. *Pivot, 3*, 28-30.

Iliams, C. (1985). Early school experiences may limit participation of women in science. In J.C. Craig and J. Hardings, (Eds.), *Contributions to the third GASAT conference* (pp. 79-85).

London: Chelsea College, University of London.

Johnston, S. (1984). Girls need a science education too. *The Australian Science Teachers Journal, 30*(2), 18-23.

Jorde, D., and Lea, A. (1987). The primary science project in Norway. In J.B. Kahle, J.Z. Daniels, and J. Harding (Eds.), *Proceedings of fourth GASAT conference* (pp. 66-72). West Lafayette, IN: Purdue University.

Kahle, J.B. (1985). Retention of girls in science: Case studies of secondary teachers. In J.B. Kahle (Ed.), *Women in science*. Philadelphia: The Falmer Press.

Kahle, J.B. (1987). SCORES: A project for change? *International Journal of Science Education, 9*(3), 325-333.

Kahle, J.B., and Lakes, M.K. (1983). The myth of equality in science classrooms. *Journal of Research in Science Teaching, 20*(2), 131-140.

Keller, E.F. (1986). How gender matters: Or why it's so hard for us to count past two. In J. Harding (Ed.), *Perspectives on gender and science* (pp. 168-183). London: The Falmer Press.

Kelly, A., Whyte, J., and Smail, B. (1984). *Final report of the GIST project*. Manchester: Manchester University, Department of Sociology.

Kelly, A. (1985). The construction of masculine science. *British Journal of Sociology of Education, 6*(2), 133-153.

LeBold, W.K. (1987). Women in engineering and science: An undergraduate research perspective. In L.S. Dix (Ed.), *Women: Their underrepresentation and career differentials in science and engineering* (pp. 49-98). Washington, DC: National Academy Press.

Linn, M.C., de Benedictis, T., Delucchi, K., Harris, A., and Stage, E. (1987). Gender differences in national assessment of education progress science items: What does "I don't know" really mean? *Journal of Research in Science Teaching, 24*(3), 267-278.

Malcolm, S.M. (1983, July). *An assessment of programs that facilitate increased access and achievement of females and minorities in K-12 mathematics and science education*. Washington, DC: American Association for the Advancement of Science, Office of Opportunities in Science.

Matyas, M.L. (1986, May). *Persistence in science-oriented majors: Factors related to attrition among male and female students*. Paper presented at American Educational Research Association, San Francisco, CA.

Mullins, I., and Jenkins, L. (1988). *The science report card*. Elements of risk and recovery (Report no. 17-S-01). Princeton: Educational Testing Service.

Newton, P. (1986). Female engineer: Femininity redefined? In J. Harding (Ed.), *Perspectives on gender and science* (pp. 40-61). London: The Falmer Press.

Parker, L.H. (1985, Dec.). *Non-sexist science education: An issue of primary concern*. Paper presented at the Science Teachers Association of Victoria Conference, Monash University, Melbourne, Australia.

Parker, L.H., and Rennie, L.J. (1986). Sex-stereotyped attitudes about science: Can they be changed? *European Journal of Science Education, 8*(2), 173-183.

Rennie, L.J., Parker, L.H., and Hutchinson, P.E. (1985). *The effect of inservice training on teacher attitudes and primary school science classroom climates* (Research Report No. 12). Perth: University of Western Australia.

Rowell, J.A. (1971). Sex differences in achievement in science and the expectations of teachers. *The Australian Journal of Education, 15*(1), 16-29.

Sadker, D., and Sadker, M. (1985). Is the ok classroom, ok? *Phi Delta Kappan, 66*(5), 358-361.

Sjoberg, S. (1986). *Naturfag og Norsk Skole* (trans. Science and the Norwegian school). The National Report of the Second International Educational Assessment in Science Study. Oslo: University of Oslo.

Smail, B. (1984). *Girl-friendly science: Avoiding sex bias in the curriculum*. London: Longman.

Smail, B. (1985). An attempt to move mountains: The 'girls into science and technology' (GIST) project. *Journal of Curriculum Studies, 17*(3), 351-354.

Spear, M.G. (1984). Sex bias in science teachers' ratings of work and pupil characteristics. *European Journal of Science Education, 6*(4), 369-377.

Tobin, K. (1987). Gender differences in science: They don't happen here! In B. Fraser and G.J. Giddings (Eds.), *Gender issues in science education* (Monograph No. 8) (pp. 37-45). Perth: Curtin University of Technology, Faculty of Education.

Tobin, K., and Gallagher, J.J. (1987). The role of target students in the science classroom. *Journal of Research in Science Teaching, 24*, 61-75.

Tobin, K., and Garnett, P. (1987). Gender differences in science activities. *Science Education, 71*(1), 91-105.

Trowbridge, L., and Bybee, R. (1987). *Becoming a secondary school science teacher*, (4th ed.). Columbus, OH: Merrill Publishing Co.

Vetter, B.M. (1987). Women's progress. *Mosaic, 18*(1), 2-9.

Ware, N.C., and Lee, V. (1987). *Sex differences in the choice of college science*. Unpublished paper, Radcliffe College, Cambridge, MA.

Weiss, I.R. (1978). Report of the 1977 National survey of science, mathematics, and social studies education (Report SE78-72). Washington, DC: U.S. Government Printing Office.

Whyte, J. (1986). *Girls into science and technology*. London: Routledge and Kegan Paul.

STUDENTS LEARN BY PRACTICE---PROBLEM SOLVING, DESIGNING AND CARRYING OUT EXPERIMENTS

The following two articles are representative of current educational research that focuses on the role of thinking and cognitive skills in learning. Both Resnick and Roth take a broader view of "practice" than previous researchers. They assert that science instruction should involve more than simply having students "do" science. Both authors emphasize that successful learning is a complex event that involves more than manipulating materials or executing routines.

Resnick comprehensively discusses educational practice in terms of cognitive skills. Her essential argument is that schools need to cultivate thinking skills and yet keep in mind that "thinking is driven by and supported by knowledge, in the form of both specific facts and organizing principles." Resnick states that students need to know both problem solving techniques and when to apply them. Roth's article focuses specifically on science education in the elementary school. In her discussion of how science should be taught, Roth considers and evaluates three perspectives in science teaching---the inquiry perspective, the Science-Technology-Society (STS) perspective, and the conceptual change perspective. She finds that the conceptual change perspective offers the most useful framework for science instruction because it encourages students to use processes (in this case, processes such as "observing" and "hypothesizing) to develop better understandings of natural phenomena. For Roth, the primary educational goal is "instruction that involves students in using these scientific processes to change their own theories in ways that are personally meaningful and consistent with scientific explanations."

EDUCATION AND LEARNING TO THINK

Lauren B. Resnick

Introduction

The question of whether schools can do a better job of teaching American children "higher order skills" is very much in the air. It arises in Congressional hearings, where calls are heard for school graduates better able to take on work that requires responsibility and judgement. It is reflected in public concern that changing employment demands are not being met, students' preparation for college is less than satisfactory, and general problem-solving abilities remain low. Yet beyond the agreement that our schools ought to be doing better than they are at building the intellectual capabilities of American young people, it is extremely difficult to discern what really should and can be done.

The first difficulties arise with the very question of what is meant by the term "higher order skills." Many candidate definitions are available. Philosophers promote critical thinking and logical reasoning skills, developmental psychologists point to metacognition, and cognitive scientists study cognitive strategies and heuristics. Educators advocate training in study skills and problem solving. How should we make sense of these many labels? Do critical thinking, metacognition, cognitive strategies, and study skills refer to the same kinds of capabilities? And how are they related to the problem solving abilities that mathematicians, scientists, and engineers try to teach their students? Are intelligence tests and scholastic aptitude tests good indicators of higher order skills, and if so, should we be teaching students the kinds of things that appear on these tests? What about artistic creativity and interpretive skill, and the ability to find and refine problems as well as to solve those others have set? And, perhaps most troubling of all, do any of these "intellectuals' concerns" really have much to do with what the vast majority of students will do in their work and personal lives after school? Do the higher order skills needed on the job or in the exercise of one's rights and duties as a citizen really depend on the kinds of abilities educators and the academic community are discussing?

---

*Reprinted with permission from **Education and Learning to Think**, 1987. Published by National Academy Press, Washington, DC.*

Mingled with the difficulty of defining higher order thinking skills is the troubling sense that there may, in fact, be little new to say about the topic. Inevitably, we hear the question: Is there really anything new about schools' trying to teach higher order skills? Haven't schools always hoped to teach students to think critically, to reason, to solve problems, to interpret, to refine ideas and to apply them in creative ways? Most of us can remember a teacher who inspired us personally in these directions, and schools everywhere include such aspirations in their statements of goals. Nevertheless, we seem to agree that students do not adequately learn these higher order abilities. Perhaps the fact that our schools have been less than successful at meeting these goals means that we have simply given up the old truths in education. Perhaps if we went back to old-fashioned courses and old-fashioned methods, the problem of teaching higher order skills would be solved without further special attention. Or, more pessimistically, perhaps we should conclude that decades of trying unsuccessfully to teach higher order skills in school show that such goals are not reachable; perhaps higher order abilities develop elsewhere than in school, and it would be wisest for schools to concentrate on the "basics," letting higher order abilities develop elsewhere than in school, and it would be wisest for schools to concentrate on the "basics," letting higher order abilities emerge later or under other auspices. To consider these fundamental questions, we need a working definition of higher order sills and an understanding of their historical role in American schools.

Higher Order Skills: A Working Definition and a Historical Perspective

Thinking skills resist the precise forms of definition we have come to associate with the setting of specified objectives for schooling. Nevertheless, it is relatively easy to list some key features of higher order thinking. When we do this, we become aware that, although we cannot define it exactly, we can recognize higher order thinking when it occurs. Consider the following:

> Higher order thinking is nonalgorithmic. That is, the path of action is not fully specified in advance.
>
> Higher order thinking tends to be complex. The total path is not "visible" (mentally speaking) from any single vantage point.
>
> Higher order thinking often yields multiple solutions, each with costs and benefits, rather than unique solutions.
>
> Higher order thinking involves nuanced judgement and interpretation
>
> Higher order thinking involves the application of multiple criteria, which sometimes conflict with one another.
>
> Higher order thinking often involves uncertainty. Not everything that bears on the task at hand is known.
>
> Higher order thinking involves self-regulation of the thinking process. We do not recognize higher order thinking in an individual when someone else "calls the plays" at every step.
>
> Higher order thinking involves imposing meaning, finding structure in apparent disorder.

Higher order thinking is effortful. There is considerable mental work involved in the kinds of elaborations and judgments required.

This broad characterization of higher order thinking points to a historical fact that is often overlooked when considering the school curriculum, a fact that helps to resolve the question of what is new about our current concerns. American schools, like public schools in other industrialized countries, have inherited two quite distinct educational traditions—one concerned with elite education, the other concerned with mass education. These traditions conceived of schooling differently, had different clienteles, and held different goals for their students. Only in the last sixty years or so have the two traditions merged, at least to the extent that most students now attend comprehensive schools in which several educational programs and student groups coexist. Yet a case can be made that the continuing and as yet unresolved tension between the goals and methods of elite and mass education produces our current concern regarding the teaching of higher order skills.

If we examine the educational institutions aimed at the elite in the population, today's higher order goals are nothing new. They represent what might be called the "high literacy" strand in the history of education (Resnick and Resnick, 1977). Since there have been books and writing, there also have been schools and related institutions established to train an intellectual elite, drawn largely from privileged social strata, in capabilities of reasoning, rhetoric, mathematical and scientific thought, and other skills that today carry the high order label. These were state, private, and religious institutions with, over the centuries, extremely varied ideas of how to go about the educational task. All were highly selective institutions. A minority of the population attended them, and this minority was selected at least in part on the basis of a taste for academic learning and the ability to perform well in a very special kind of institution.

In America, various "academies," some private and some public, carried on this tradition through the nineteenth century and into the twentieth. Until they began to be transformed early in this century, even public high schools were in the academy mold. Only a minority of young people attended or even thought of attending them. There were entrance examinations. The curriculum was quite strictly academic. Extensive writing, textual criticism, and the like were expected. Although today we might not recognize nineteenth-century curricula as promoting creative thinking or independent problem solving, the elite academies expected to produce, and to a considerable extent succeeded in producing, intellectual performance beyond the ordinary.

Historically, it must be stressed, the academies did not treat education of the full population of young people as within their purview. Schools for the masses arose from different roots and are a much more recent phenomenon in the history of education. Mass education derives from a "low literacy" tradition (Resnick and Resnick, 1977) aimed at producing minimal levels of competence in the general population. It originated in Europe in Reformation and counter-Reformation efforts to produce a literate, catechism- and bible-reading population. During the nineteenth century, mass schooling was adopted as part of a new national agenda in countries that were just beginning to form citizen armies and to impose common language and culture on their populations. In the United States, village and township schools were established early, probably reflecting radical Protestant traditions as well as new definitions of citizenship appropriate to the new nation. Throughout the nineteenth century, this nation knew levels of school attendance and literacy ahead of most other countries, despite the continuing flow of poor and poorly educated immigrants. As cities began to grow, massive urban school systems grew as well. Only racial minorities were systematically excluded or separated within the schools.

The mass education system that evolved under these circumstances focused largely on elementary schooling, and rather sharp divisions between elementary and secondary education persisted. This distinction was apparent both in who went to school and in what was taught. Almost everyone went to elementary school, although a limited number finished the entire eight-year course. Only a few went to high school or its equivalent. The elementary schools served the masses and concerned themselves with basic skills of reading and computation, with health and citizenship training, and the like. Routinized performance rather than creative and independent thought was stressed. Mass education was, form its inception, concerned with inculcating routine abilities: simple computation, reading predictable texts, reciting religious or civic codes. It did not take as goals for its students the ability to interpret unfamiliar texts, create material others would want and need to read, construct convincing arguments, develop original solutions to technical or social problems. The political conditions under which mass education developed encouraged instead the routinization of basic skills as well as the standardization of teaching and education institutions. Standardization was a means of ensuring that at least minimal curriculum standards would be met, that teachers would be hired on the basis of competency for the job rather than political or familial affiliation, and that those responsible for the expenditure of public funds could exercise orderly oversight over the educational process. Standardized testing was one of the methods developed to exercise oversight and centralized control of the schools (Resnick, 1980).

Early in the twentieth century, the institutional division between routine-oriented elementary schools and secondary academies in the high literacy tradition began to dissolve. Responding to changing economic and social conditions, more and more young people began to seek high school education, and educators gradually began to treat secondary education of a much larger and more varied population as being their proper concern. The secondary schools were over the next decades to become the mass institutions the elementary schools had been. The growth of this new secondary school population marked the beginning of the debate that continues even today. This debate concerns what the appropriate curriculum ought to be for secondary schools designed to serve everyone. The terms of the debate were set, in great part, by a National Education Association (NEA) commission report entitled The Cardinal Principles of Secondary Education (Bureau of Education, 1918). The report provided a theory and ideology for the place of a vocationally oriented curriculum in high school as part of a diversified secondary program adapted to different types of students. This represented a clear challenge to the older ideology that organized the high school curriculum around a common core of the traditional liberal disciplines.

The tension between vocationalism and traditional disciplines as the center of the high school program has never been resolved. Responding to post-World War II manpower needs, the 1950s and early 1960s saw a greater emphasis on traditional disciplines, especially mathematics and science. Yet political and social pressures from many quarters sustained the demand for vocational training and other programs designed to keep students in school as long as possible. Other developments in the later 1960s and 1970s led to a near-complete abandonment of the traditional core curriculum, even for students who had been its traditional consumers. Schools continued to require academic courses, but the requirements were often minimal and course content focused increasingly on application and practical topics—often replacing more traditional, demanding material. Written composition and other activities that engaged higher order skills all but disappeared from the curriculum.

The effect of all this has been to reduce, and sometimes to drive out of existence, the high literacy goals that had been the focus of the academies and their preparatory institutions. Yet the taste for such goals has survived and can be seen in recent efforts to revive interest in higher order skills teaching. This revival, however, takes place in an educational and social context that dictates an extension of high literacy goals to a much broader segment of the population than has ever before been considered capable of such learning. Today, we are committed to educating all Americans in the secondary schools and a large proportion (higher than in any other country in the world) in some form of postsecondary institution. These students' educational needs cannot be met by traditional vocational programs that no longer prepare students for productive participation in an increasingly diversified economic environment. Employers today complain that they cannot count on schools and colleges to produce young people who can move easily into more complex kinds of work. They seem to be seeking general skills such as the ability to write and speak effectively, the ability to learn easily on the job, the ability to use quantitative skills needed to apply various tools of production and management, the ability to read complex material, and the ability to build and elevate arguments. These abilities go well beyond the routinized skills of the old mass curriculum. In fact, they are much like the abilities demanded for college-bound students in the College Board's book, Academic Preparation for College (College Entrance Examination Board, 1983). Yet teaching such competencies to the mass of students remains a considerable challenge.

This, then, is part of what is new about the current drive for teaching higher order skills. The goals of increasing thinking and reasoning ability are old ones for educators. Such abilities have been the goals of some schools at least since the time of Plato. But these goals were part of the high literacy tradition; they did not, by and large, apply to the more recent schools for the masses. Although it is not new to include thinking, problem solving, and reasoning in someone's school curriculum, it is new to include it in everyone's curriculum. It is new to take seriously the aspiration of making thinking and problem solving a regular part of a school program for all of the population, even minorities, even non-English speakers, even the poor. It is a new challenge to develop educational programs that assume that all individuals, not just an elite, can become competent thinkers.

The Nature of Thinking and Learning: Going Beyond the Routine

This challenge comes at a time when we also have new knowledge about the nature of thinking and strong hints about how thinking abilities are learned. In the last decade or two, cognitive science research has allowed us to look into the thinking mind, figuratively at least, and to specify more precisely the reasoning processes of both successful and less successful thinkers (Newell and Estes, 1983). More recently, researcher have begun to investigate how the ability and the propensity to think well are acquired and maintained. These two bodies of research—on the nature of human thinking and on the acquisition of thinking and learning skills—are beginning to make explicit what we mean by higher order skills and what means of cultivating such skills are most likely to be successful. This process of making explicit the abilities formerly left to the intuitions of gifted learners and teachers is precisely what we need to establish a scientific foundation for the new agenda of extending thinking and reasoning abilities to all segments of the population.

The most important single message of modern research on the nature of thinking is that the kinds of activities traditionally associated with thinking are not limited to advanced levels of development. Instead, these activities are an intimate part of even elementary levels of reading, mathematics, and other branches of learning—when learning is proceeding well. In fact, the term "higher order" skills is probably itself fundamentally misleading, for it suggests that another set of skills, presumably called "lower order," needs to come first. This assumption—that there is a sequence from lower level activities that do not require much independent thinking or judgement to higher level ones that do—colors much of educational theory and practice. Implicitly at least, it justifies long years of drill on the "basics" before thinking and problem solving are demanded. Cognitive research on the nature of basic skills such as reading and mathematics provides a fundamental challenge to this assumption. Indeed, research suggests that failure to cultivate aspects of thinking such as those listed in our working definition of higher order skills may be a source of major learning difficulties even in elementary school.

[Resnick here examines reading as a higher order skill and meaning construction in mathematics.]

General Reasoning: Improving Intelligence

Mathematics and reading are not unique in the extent to which high-level performance depends on processes of monitoring one's understanding, imposing meaning and structure, and raising questions about presented material. Much the same story can be told about all the subject matter in the school curriculum and about all but the most routine job performances. Recent research in science problem solving, for example, shows that experts do not respond to problems as they are presented—writing equations for every relationship described and then using routine procedures for manipulating equations. Instead, they reinterpret the problems, recasting them in terms of general scientific principles until the solutions become almost self-evident (Larkin et al., 1980). Expert writers treat the process of composing an essay as a complex task of shaping a communication that will appeal to and convince an intended audience rather than as a simple task of writing down everything they know about a topic (Bereiter and Scardamalia, 1982; Flower and Hayes, 1980). In the social sciences, trained thinkers call upon a wide range of knowledge relevant to a topic to construct proposals for action and to build justifications for those proposals that conform to many of the classical principles of rhetorical argumentation (Voss et al., 1983). Skilled technicians repairing equipment do not just proceed through routine checklists; instead, they construct "mental models" of complex systems and use them to reason about possible causes of observed breakdowns and potential repairs (e.g., de Kleer and Brown, 1980).

In all of these cases, certain kinds of higher order thinking recur: experts elaborate and reconstruct problems into new forms; they look for consistencies and inconsistencies in proposed solutions; they pursue implications of initial ideas and make modifications rather than seeking quick solutions and sticking with initial ideas; they reason by analogy to other, similar situations. These similarities, long noted in discussions of intelligence (see Journal of Educational Psychology, 1921; Simon, 1976; Sternberg and Detterman, 1979) and problem solving (Tuma and Reif, 1980), lead naturally to the question of whether there might not be some general thinking skills that would produce improved ability to learn across many traditional curriculum areas. If such skills exist and if we can find effective ways to teach them, we can imagine an important increase in educational efficiency, for—it would seem—a relatively narrow instructional effort might produce wide learning results.

The search for general learning skills is not a new one. Both educators and psychologists have long sought to identify and characterize such skills, the former because of the educational efficiency such skills could help them realize, the latter in search of unifying characteristics of human thought. Psychological research gives us reason to believe in the reality of general skills for learning as well as reason to maintain a degree of skepticism. In the next section we will review recent efforts to teach higher order skills. These efforts provide the newest body of evidence on the question of whether such skills are teachable. Before proceeding to the review, however, we should first consider what the body of past research would suggest.

Past Research

Psychometric research provides the best-established evidence for the existence of cognitive skills that play a role in diverse kinds of learning. When two or more cognitive abilities are tested, there is almost always a positive correlation between the measures. People who do well on one ability test are, on the average, likely to do well on the others. Virtually the only conditions under which such a correlation is not found are those in which tests have been specifically designed not to correlate. For example, investigators have built tests of creativity explicitly designed to be psychometrically independent of IQ. Tests that correlate positively are presumed to share underlying processes. The fact that most intelligence do correlate strongly and that a general factor can always be identified through statistical methods such as factor analysis suggests that all tests have some processes in common. These common processes, are presumably, general abilities.

Of course, such findings only raise new questions. How do we characterize these common processes? Is there reason to think they are teachable? When cognitive scientists do information-processing analyses of complex skills, they find the same kinds of basic problem-solving processes used in task after task (Simon, 1976). For example, one of the earliest uses of computers to explore processes of human reasoning resulted in the construction of a program that solved symbolic logic problems. This program was called the General Problem Solver (GPS) in the belief that its processes would play a role in solving many kinds of problems, not just those of symbolic logic. This has turned out to be partly true. Although GPS itself can solve only a limited range of problems, the kinds of processes used by GPS appear over and over again in simulations of human performance of complex tasks. Processes such as means-ends analysis (comparing one's final goal with results that would be produced by procedures currently available), subgoal formation (forming a new goal that is easier to solve and that is en route to the final goal), generate-and-test routines (generating actions and testing them against constraints), and other general problem-solving routines are used in tasks as varied as buggy arithmetic routines, planning compositions, constructing geometry proofs, and troubleshooting electronic devices. The reason that a single artificial intelligence program cannot solve a wide variety of problems is not that the fundamental processes it applies are widely different across domains, but rather that the program must apply these processes to very specific, organized bodies of knowledge, and so it becomes specific to its knowledge base (see Dehn and Schank, 1982).

Other processes that appear repeatedly in analyses of complex task performance play a kind of "executive" or self-regulatory role in thinking. People use these processes to keep track of their own understanding, to initiate review or rehearsal activities when needed, and to deliberately organize their attention and other resources in order to learn something. These activities have been shown to be characteristic of effective learners, good readers and writers, and strong problem solvers. The same processes are relatively absent in younger or less intelligent individuals. These skills are sometimes called "metacognitive skills" (see Brown et al., 1983) because they operate on an individual's own cognitive processes. They have been suggested frequently as processes that could be taught and that would enhance learning and thinking in a wide range of specific situations.

The problem-solving skills identified in cognitive simulation research and the metacognitive skills identified in developmental psychology research have both been proposed as candidates for teaching. The hope is held out that if we can improve specific skills through some form of direct teaching, then people's ability to perform various kinds of learning, thinking, and problem-solving tasks in which such skills have been observed will also improve.

On the other hand, the very body of research that has helped to identify the candidate "general" skills also provides reason for questioning their educational importance. Cognitive research yields repeated demonstrations that specific content area knowledge plays a central role in reasoning, thinking, and learning of all kinds. I have already alluded to several examples of the importance of specific knowledge. Specific knowledge about a text's topic affects processes of language comprehension, for example. Skilled science problem solvers rely on their knowledge of scientific principles to recast problems into more elegant and easily solvable forms. Political scientists' argumentation becomes degraded when they know little about the particular problem or the particular part of the world under discussion (Voss et al., 1983). Even on the tasks used to assess general intelligence or scholastic aptitude, recent analyses have made it clear that much depends on specific knowledge: of vocabulary, of particular number relationships, of possible transformations of visual displays, and the like (cf. Glaser, 1984). General skills such as breaking down a problem into simpler problems or checking to see whether one has captured the main idea of a passage may be impossible to apply if one does not have a store of knowledge about similar problems—or know enough about the topic to be able to recognize its central ideas. Of course, to appreciate the dependence of general skills application on specific knowledge is not to deny that such general skills exist. Yet such an understanding raises questions about the wisdom of attempting to develop thinking skills outside of the context of specific knowledge domains. It suggests that a more promising route may be to teach thinking skills within specific disciplines and perhaps hope for some transfer to other disciplines as relevant knowledge is acquired.

On first consideration the hope for transfer of thinking abilities across disciplines seems misplaced. A long history of research exists on transfer among school subjects. Over the decades, educators have espoused a recurring belief that certain school subject matters "discipline the mind" and therefore should be taught not so much for their inherent value as for their efficacy in facilitating other learning. Latin was defended for many years in these terms; mathematics and logic are so defended today. Most recently, computer programming has been proposed as a way to develop general problem-solving and reasoning abilities (e.g., Papert, 1980). The view that we can expect strong transfer from learning in one area to improvements across the board has never been well supported empirically. At the turn of the century, Thorndike and Woodworth (1901) studied transfer among school subjects and found that it was more efficient to study the subject of interest (English vocabulary, for example) than to study some other subject (e.g., Latin) that "prepared" one's mind. Subsequent reviews of research on transfer of school subject matter generally have confirmed Thorndike and Woodworth's finding.

Nevertheless, the history of transfer research need not be totally discouraging; most of this research does not directly address the questions of most concern to those whose goal is the improvement of general thinking and learning abilities. First, the subject matter teaching in these studies has rarely been aimed at developing transferable skill and knowledge. We thus do not know what leverage there might be in instruction explicitly aimed at producing general skills in the context of a particular discipline. Second, evaluations of learning outcomes have rested mainly on what knowledge was acquired in the transfer discipline, rather than on whether skills for acquiring knowledge in that discipline have been enhanced. The issue of transferability of thinking and learning skills, then, is still open.

[Resnick here discusses current programs for teaching higher order skills and some related problems of assessment. She then examines thinking in terms of the school curriculum.]

Cultivating the Disposition to Higher Order Thinking

It has been convenient to examine teaching programs in several distinct categories. Yet there are striking points of similarity among those programs that have shown some promising results. Many such programs rely on a social setting and social interaction for much of teaching and practice. Although one can imagine individually worked exercises designed to improve aspects of thinking skill, very few programs in fact propose such activities. Instead, students are encouraged to work problems in pairs or in small groups. Instructors may also orchestrate special discussion and practice sessions. When investigators of different theoretical orientations and disciplinary backgrounds converge on a common prescription in this way, we should consider what shared intuition may be at work. What roles might social interaction be playing in the development of thinking? The authors cited in the preceding pages mention several possibilities.

First, the social setting provides occasions for modeling effective thinking strategies. Skilled thinkers (often the instructor but sometimes more advanced fellow students) can demonstrate desirable ways of attacking problems, analyzing texts, and constructing arguments. This process opens normally hidden mental activities to inspection. Through observing others, students can become aware of mental processes that might otherwise have remained entirely implicit. Research suggests, however, that modeling alone does not produce very powerful results. If students only watched more skilled thinkers perform, they would not substantially improve their own thinking.

Apparently there is more to learning in a social setting than watching other perform. "Thinking aloud" in a social setting allows others—peers or an instructor—to critique and shape one's performance, something that cannot be done effectively if only the results but not the thought processes are visible. The social setting may also provide a kind of scaffolding for an individual learner's initially limited performance. Instead of practicing small bits of thinking in isolation with no sense of each bit's significance to the task as a whole, a group solves a problem, or writes a composition, or analyzes an argument together. Within the group, extreme novices can participate in performing complex tasks. If things go well, they can eventually take over most or all of the work themselves, with a developed appreciation of how individual elements in the process contribute to the whole. This theory, adapted from Vygotsky (1978), is embodied explicitly in the reciprocal teaching of Palincsar and Brown, and variants of it have been proposed by a number of other investigators (e.g., Collins et al., in press).

The social setting may also function to motivate students. Students are encouraged to try new, more active approaches, and they receive social support even for partially successful efforts. Through this process, students come to think of themselves as capable of engaging in independent thinking and of exercising control over their learning processes. The public setting also lends social status and validation to what can perhaps best be called the disposition to higher order thinking. The term disposition should not be taken to imply a biological or inherited trait. As used here, it is more akin to a habit of thought, one that can be learned and, therefore, taught. Engaging in higher order thinking with others seems likely to teach students that they have the ability, the permission, and even the obligation to engage in a kind of critical analysis that does not always accept problem formulations as presented or that may challenge an accepted position.

We have good reason to believe that shaping this disposition to critical thought is central to developing higher order cognitive abilities in students. Research on strategy training shows that, if instruction is to work at all, it often works very quickly—in just a few lessons or sometimes with little more that directions to use some strategy. However, people induced to use a particular learning strategy will often do so on the immediate occasion but will fail to apply the same strategy on subsequent occasions. Both of these recurrent findings serve to remind us that much of learning to be a good thinker is learning to recognize and even search for opportunities to apply one's mental capacities (cf. Belmont et al., 1982).

This suggests that the task for those who would raise the intellectual performance levels in children is not just to teach children new cognitive processes but to get them to use those processes widely and frequently. The kind of higher order thinking we have discussed requires elaborating, adding complexity, and going beyond the given to construct new formulations of issues. It also involves weighing multiple alternatives and sometimes accepting uncertainty. As such, higher order thinking requires effort on the part of the individual and may involve some social risk—of disagreeing with others perceived to be more powerful, of not arriving at the expected answers, of not always responding instantly. To overcome these difficulties, educational institutions must cultivate not only skills for thinking but also the disposition to use them.

A widely shared set of implicit assumptions exists about how dispositions for higher order thinking might develop. They center on the role of a social community in establishing norms of behavior, providing opportunity for practice, and providing occasions for learning particular skills. The fundamental theme is that such dispositions are cultivated by participation in social communities that value thinking and independent judgement. Such communities communicate these values by making available many occasions for such activity and responding encouragingly to expressions of questioning and judgement. The process of learning is further aided when there are many opportunities to observe others engaging in such thinking activities. Finally, dispositions for higher order thinking require sustained long-term cultivation; they do not emerge from short-term, quick-fix interventions.

This set of beliefs, although highly plausible, has received little empirical investigation. On the whole, research on the development of cognitive abilities has proceeded quite separately from research on social and personality development. For example, the extensive body of childhood socialization research (Hetherington, 1983) says much about the emergence of traits such as aggressiveness, dependency, conformity, or gender identification, but it says little about how intellectual tendencies develop. An interesting new research project (Caplan, 1985) on the development of intellectual curiosity in young children appears to be a first link between research on child socialization and our present concern for shaping higher order thinking dispositions.

"Cognitive styles" (e.g., Messick, 1976) such as reflectivity are known to be related to school performance, and efforts have been made to shape reflectivity (e.g., Meichenbaum, 1985). But this research has not generally attended to the qualitative aspects of intellectual performance, and it is impossible to know whether higher order thinking was in fact improved. Other research on improving persistence (e.g., Turkewitz et al., 1975) has tended to measure how much work students do but not whether they engage in complex cognitive activities. some recent research on intrinsic motivation may help tie motivation to the quality as well as the quantity of educational work (see Lepper, 1981, 1983; Nicholls, 1983). When people work to gain praise, grades, or material benefits, they are externally motivated. When they work to master a task, they are intrinsically motivated. Apparently some correlation exists between the kinds of motivations that keep people working and several qualitative features of their work: for example, the complexity of tasks they choose to work on, the range of material to which they attend, and the extent to which they are able to shift direction ("break set") to pursue a new, more fruitful approach (Condry and Chambers, 1981; Kruglanski, 1981; Lepper and Greene, 1981; McGraw, 1981).

A promising link between quality of thinking and persistence is being forged by investigators studying differences in people's conceptions of ability. For example, Dweck and her colleagues (Dweck, in press; Dweck and Elliot, 1983) have shown that individuals differ fundamentally in their conceptions of intelligence and that these conceptions mediate very different ways of attacking problems. A distinction is made between two competing conceptions of ability, or "theories of intelligence," that people may hold. One, called entity conception, treats ability as a global stable quality. The second, called the incremental conception, treats ability as a repertoire of skills that can be expanded through efforts to learn. Entity conceptions orient children toward performing well so that they can display their intelligence and toward not revealing lack of ability by giving "wrong" responses. Incremental conceptions orient children toward learning well so that they can acquire new knowledge or skill. Most relevant to the present argument, incremental conceptions of ability and associated learning goals lead children to analyze tasks and to formulate strategies for overcoming difficulties. We can easily recognize these as close cousins to the kinds of higher order thinking discussed in this essay. In a related analysis, Covington (1983) suggests that people who view ability as created through strategic self-management (of study time, of types of elaboration, of ways of attacking tasks) will be better able to compensate for self-attributions of low initial ability.

A key question, of course, is whether these differences in type of motivation or theory of intelligence can be deliberately shaped by the way in which school activity is organized. Evidence suggests that the nature of the environment in which one works makes a difference in whether one invokes internal or external motivations for one's work. However, research has not examined whether personal traits favoring internal motivation can be developed by deliberately altering institutional or social patterns. Very recent work by Dweck and her colleagues is examining ways of helping students to acquire and apply incremental conceptions of intelligence, but more extensive research is required before clear conclusions can be drawn. In any case, these lines of motivation research highlight the possibilities for an important convergence between efforts aimed at teaching higher order cognitive skills and those aimed at cultivating dispositions to apply those skills.

Summary and Conclusions

What Are Higher Order Skills?

Higher order thinking is difficult to define but easy to recognize when it occurs.

Higher order thinking involves a cluster of elaborative mental activities requiring nuanced judgment and analysis of complex situations according to multiple criteria. Higher order thinking is effortful and depends on self-regulation. The path of action or correct answers are not specified in advance. The thinker's task is to construct meaning and impose structure on situations rather than to construct meaning and impose structure on situations rather than to expect to find them already apparent.

Higher order thinking has always been a major goal of elite educational institutions. The current challenge is to find ways to teach higher order thinking within institutions committed to educating the entire population.

In its origins, the mass educational system was concerned with routine competencies such as simple computation, reading familiar and predictable texts, and acquiring well-defined vocational competencies. It was not considered necessary or possible for all students to learn to interpret complex texts, write extended arguments, or develop original solutions to problems. However, changing economic and social conditions are now creating a demand for these abilities in all citizens, and schools are seeking ways to cultivate thinking skills in all students. No educational system has ever been built on the assumption that everyone, not just an elite, can become a competent thinker. We must view this new challenge as an invitation to inventive and very demanding educational reform.

Higher order thinking is the hallmark of successful learning at all levels—not only the more advanced.

The challenge to reform comes at a time when cognitive research provides an important reconceptualization of the nature of thinking and learning that can inform and guide educational work. The most important single message of this body of research is that complex thinking processes—elaborating the given material, making inferences beyond what is explicitly presented, building adequate representations, analyzing and constructing relationships—are involved in even the most apparently elementary mental activities. Children cannot understand what they read without making inferences and using information that goes beyond what is written in the text. They cannot become good writers without engaging in complex problem-solving-like processes. Basic mathematics will not be effectively learned if children only try to memorize rules for manipulating written numerical symbols. All of this implies that "basic" and "higher order" skills cannot be clearly separated.

Good thinking depends on specific knowledge, but many aspects of powerful thinking are shared across disciplines and situations.

A central issue, both for educational practice and for research that can guide that practice, is whether thinking and learning abilities are general—that is, applicable in all domains of thinking—or specific to a particular domain. The evidence shows clearly that thinking is driven by and supported by knowledge, in the form of both specific facts and organizing principles. This knowledge, together with the automated recognition and performance that come with extended practice, allows experts in any field to engage in more sophisticated thinking than people new to the field. At the same time, many aspects of thinking are shared across fields of expertise. These include a wide range of oral and written communication skills, mathematization and representational abilities, principles of reasoning, and skills of argument construction and evaluation. These can be thought of as "enabling skills" for learning and thinking. Generally speaking, people rely on powerful but only narrowly applicable thinking methods in domains in which they are expert and use broadly applicable but weak methods for learning and thinking in fields they know little about. Good thinkers need both the powerful but specific and the general but weak kinds of skills.

Can Higher Order Thinking Be Directly Taught?

Elements of thinking are clearly teachable.

The programs reviewed here show that many components of thinking can be effectively taught. That is, there is evidence that the particular performances taught in the programs are in fact learned by students. The kinds of components that have been successfully taught include generating multiple ideas and alternative viewpoints on a particular topic, generating multiple ideas and alternative viewpoints on a particular topic, generating summaries, skimming, figuring out word meanings from context, solving analogies and logical puzzles, and detecting logical reasoning fallacies.

However, an integrated ability to learn, think, and reason and a broad disposition to engage in higher order thinking are not necessarily ensured by acquiring particular components of thinking.

We need direct assessments of the kinds of complex reasoning and problem-solving skills that constitute higher order thinking. Most evaluations have not made such assessments. They have relied instead on assessments of particular elements that are taught or on "indicator" tests—such as IQ or SAT scores—that are normally correlated with successful learning and thinking. However, under changed instruction and learning conditions, these traditional indicators may no longer be valid. Thus, we have less evidence than would be desirable, and less than the proliferation of programs would suggest, on whether and how thinking abilities that are integrated and usable can actually be cultivated.

Only a few programs provide convincing evidence that broadly applicable and integrated abilities have been acquired.

In the most convincing cases, improvements due to instruction have been demonstrated for reading comprehension, general grade averages, and essay writing. Some programs also demonstrate improved problem-solving or laboratory performance in specific disciplines, especially in mathematics and science, thus meeting their own goals—although not demonstrating (and not necessarily seeking) transfer to other disciplines or to practical life. A larger number of programs point to student claims that they now use the kinds of abilities taught. However, these claims are difficult to evaluate; they show that students generally feel better about their thinking and learning abilities after the course, but they do not tell us whether these improved self-assessments are in fact warranted.

Current testing practices in American education do not provide very powerful tools for assessing the effects of efforts to teach thinking and reasoning. Testing practices may in fact interfere with cultivation of the kind of higher order skills desired.

In general, the tests used in assessing educational efforts involve multiple choice or other short, precoded answers. These tests can measure the accumulation of knowledge and can be used to examine specific components of reasoning or thinking. However, they are ill suited to assessing the kinds of integrated thinking that we call "higher order." If progress is to be made in converting American schools to the higher order thinking agenda, we must develop forms of assessment that are more suited to the nature of the abilities we seek to teach.

How Should Instruction in Higher Order Thinking Be Organized?

A broad disposition to higher order thinking must be cultivated.

Isolated instruction in thinking skills, no matter how elegant the training provided, is unlikely to produce broadly used thinking ability. Thinking well requires more than knowing a selected set of strategies or techniques for problem solving and learning. It also requires the motivation to apply them, even though they may involve more effort than routine performances as well as some risk of social controversy. This implies that higher order skills must suffuse the school program from kindergarten on and in every subject matter. Training in general skills must be supplemented and supported by application throughout the curriculum. Various subject matters in the school program should be taught with an eye to developing the powerful thinking methods used by experts in those disciplines. Students must come to think of themselves as able and obligated to engage in critical analysis and problem solving throughout schooling. The following are promising directions that educational experimentation might take.

Embedding instruction in thinking skills within the academic disciplines of the school has several advantages.

It ensures that there is something solid to reason about. It supplies criteria from within the disciplinary traditions for what constitutes good reasoning and thinking. It ensures that something worthwhile will have been taught and learned even if wide transfer proves impossible. However, there is a caveat for those who seek to embed higher order skills teaching in the existing school program. Thinking skills tend to be driven out of the curriculum by ever-growing demands for teaching larger and larger bodies of knowledge. The idea that knowledge must be acquired first and that its application to reasoning and problem solving can be delayed is a persistent one in educational thinking. "Hierarchies" of educational objectives, although intended to promote attention to higher order skills, paradoxically feed this belief by suggesting that knowledge acquisition is a first stage in a sequence of educational goals. The relative ease of assessing people's knowledge, as opposed to their thought processes, further feeds this tendency in educational practice.

Periodically, educators resist this pressure by proposing that various forms of process- or skill-oriented teaching replace knowledge-oriented instruction. In the past, this has often led to a severe deemphasis of basic subject matter knowledge. This, in turn, has had the effect of alienating many subject matter specialists, creating pendulum swings of educational opinion in which knowledge-oriented and process-oriented programs periodically displace each other, and delaying any serious resolution of the knowledge-process paradox. We cannot allow these pendulum swings to continue. Cognitive research shows the intimate relationship of subject matter knowledge and reasoning processes. We need both practical experimentation in schools and more controlled instructional experimentation in laboratories to discover ways of incorporating our new understanding of the knowledge-reasoning connection into instruction.

Reorienting instruction in the 3-Rs (the "enabling disciplines") so that they incorporate more of the higher order processes seems a particularly promising approach to improving thinking skills.

The 3-Rs of the traditional basic school curriculum can become the environment for higher order education. Effective reading, writing, and mathematics learning depend on elaboration, explication, and various forms of meaning construction. Reorienting basic instruction in these curricula to focus on intentional, self-managed learning and strategies for meaning construction, rather than on routinized performances, will result in more effective basic skills instruction while providing a strong base for higher order skill development in other disciplines.

A fourth "R"—reasoning—might be considered a candidate for a new enabling discipline in the school curriculum.

Many philosophers argue that principles of logical reasoning are unitary and not specific to particular domains of knowledge. The study of reasoning, they claim, can enable effective thinking across the disciplines. Although there has been little empirical investigation of this claim, the hypothesis is a reasonable one and should be investigated carefully. A potential pitfall is that learning to identify reasoning fallacies—a core element of most programs in informal logic and critical thinking—may not in fact help people improve their own reasoning. This question needs careful attention, with appropriate evaluation of the extent to which students in reasoning courses learn to produce, as well as analyze, reasoned arguments.

Links between thinking skills and motivation for thinking must be developed.

Everyone agrees that successful educational achievement requires both motivation and appropriate cognitive activity. Yet our theories implicitly treat motivation and cognition as if they worked independently to determine the nature and extent of learning. In fact, these traditionally separate factors appear far more intimately related than most current research helps us to appreciate. However, recent research linking children's conceptions of their own and others' intelligence to the ways in which they analyze learning tasks offers a promising new connection, as does research on intrinsic motivation for learning. Active experimentation on what kinds of school activity organization cultivate motivation for particular kinds of complex and strategic learning is needed. The two concerns must be merged as this work proceeds; efforts to develop more intellectually functional motivational patterns should not become substitutes for efforts to establish specific cognitive competencies. Motivation for learning will be empty if substantive cognitive abilities are not developed, and the cognitive abilities will remain unused if the disposition to thinking is not developed.

References

Arbitman-Smith, R., Haywood, H.C., and Bransford, J.D. (1984). Assessing cognitive change. In P. Brooks, R. Sperber, and C.N. McCauley (Eds.), *Learning and cognition in the mentally retarded* (pp. 433-472). Hillsdale, NJ: Erlbaum.

Belmont, J.M., Butterfield, E.C., and Ferretti, R.P. (1982). To secure transfer of training, instruct self-management skills. In D.K. Detterman and R.J. Sternberg (Eds.), *How and how much can intelligence be increased?* (pp. 147-154). Norwood, NJ: Ablex.

Bereiter, C., and Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and Instruction, 2*, 131-156.

Bereiter, C., and Scardamalia, M. (1982). From conversation to composition: The role of instruction in a developmental process. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 2, pp. 1-64) Hillsdale, NJ: Erlbaum.

Bolt Beranek and Newman. (1983). *Final report, Project Intelligence: The development of procedures to enhance thinking skills.* Cambridge, MA: Bolt Beranek and Newman Laboratories, Inc.

Bransford, J.D. (In press). *Enhancing thinking and learning.* San Francisco: Freeman and Co.

Brown, A.L., Bransford, J.D., Ferrara, R.A., and Campione, J.C. (1983). Learning, remembering, and understanding. In J.H. Flavell and E.M. Markman (Eds.), *Cognitive development* (Vol. III of P.H. Mussen, Ed., Handbook of child psychology, pp. 77-166). New York: Wiley.

Brown, J.S., and Van Lehn, K. (1980). Repair theory: A generative theory of "bugs." In T.P. Carpenter, J.M. Moser, and T.A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 117-135). Hillsdale, NJ: Erlbaum.

Browne, M.N., and Keeley, S.N. (1981). *Asking the right questions.* Englewood Cliffs, NJ: Prentice-Hall.

Bureau of Education. (1918). *Cardinal principles of secondary education* (Bull. No. 35). Washington, DC: Department of the Interior.

Caplan, J.S. (1985, April). The development of intellectual curiosity. Paper presented at the meeting of the Society for Research in Child Development, Toronto, Canada.

Carraher, T.N., Carraher, D.W., and Schliemann, A.D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology, 3*, 21-29.

Clement, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception. *Journal of Research in Mathematics Education, 13*, 16-30.

Cole, M., and Griffen, P. (Eds.). (1987). *Contextual factors in education: Improving science and math education for minorities and women*. Madison, WI: Wisconsin Center for Educational Research.

College Entrance Examination Board. (1983). *Academic preparation for college*. New York: College Entrance Examination Board.

Collins, A., Gentner, D., and Rubin, A. (1981). *Teaching study strategies* (Rep. No. 4794). Cambridge, MA: Bolt Beraneck and Newman Laboratories, Inc.

Collins, A., and Smith, E.E. (1982). Teaching the process of reading comprehension. In D. Detterman and R. Sternberg (Eds.), *How much and how can intelligence be increased?* (pp. 173-185). Norwood, NJ: Ablex.

Collins, A., Brown, J.S., and Newman, S.E. (In press). Teaching the craft of reading, writing, and mathematics. In L.B. Resnick (Ed.), *Knowing and learning: Issues for a cognitive science of instruction*. Hillsdale, NJ: Erlbaum.

Condry, J., and Chambers, J. (1981). Intrinsic motivation and the process of learning. In M.R. Lepper and D. Greene (Eds.), *The hidden costs of reward* (pp. 61-84). Hillsdale, NJ: Erlbaum.

Covington, M.V. (1983). Motivated cognitions. In S.G. Paris, G.M. Olson, and H.W. Stevenson (Eds.), *Learning and motivation in the classroom* (pp. 139-164). Hillsdale, NJ: Erlbaum.

Covington, M.V. (1985). Strategic thinking and the fear of failure. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 389-416). Hillsdale, NJ: Erlbaum.

Covington, M.V. (In press). Instruction in problem-solving planning. In S.L. Friedman, E.K. Scholnick, and R.R. Cocking (Eds.), *Blueprints for thinking: The role of planning in cognitive development*. Cambridge, England: Cambridge University Press.

Dansereau, D.F. (1985). Learning strategy research. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 209-239). Hillsdale, NJ: Erlbaum.

Day, J.D. (1980). Training summarization skills: A comparison of teaching methods. Unpublished doctoral dissertation, University of Illinois.

de Bono, E. (1970). *Lateral thinking*. New York: Harper & Row.

de Bono, E. (1976). *Teaching thinking*. London: Temple Smith.

de Bono, E. (1985). The CoRT thinking program. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 363-388). Hillsdale, NJ: Erlbaum.

Dehn, N., and Schank, R. (1982). Artificial and human intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 353-391). Cambridge, MA: Harvard University Press.

de Kleer, J., and Brown, J.S. (1980). Mental models of physical mechanisms and their acquisition. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 285-309). Hillsdale, NJ: Erlbaum.

Dweck, C.S. (In press). Motivation. In R. Glaser and A. Lesgold (Eds.), *The handbook of psychology and education* (Vol. 1). Hillsdale, NJ: Erlbaum.

Dweck, C.S., and Elliot, E.S. (1983). Achievement motivation. In E.M. Hetherington (Ed.), Socialization, personality, and social development (Vol. IV of P.H. Mussen, Ed., *Handbook of child psychology*, pp. 643-692). New York: Wiley.

Edwards, J., Baldauf, R.B., and Cook, J. (1984, August). The effects of a thinking skills program on students. Paper presented at the Conference on Thinking, Harvard University, Cambridge, Massachusetts.

Ennis, R.H. (1980). A conception of rational thinking. In J. Coombs (Ed.), *Philosophy of education 1979* (pp. 3-30). Normal, IL: Philosophy of Education Society.

Ennis, R.H. (1985). Critical thinking and the curriculum. *National Forum, 65* (1), 28-31.

Feuerstein, R., Jensen, M.R., Hoffman, M.B., and Rand, Y. (1985). Instrumental enrichment, an intervention program for structural cognitive modifiability: Theory and practice. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 43-82). Hillsdale, NJ: Erlbaum.

Flower, L.S., and Hayes, J.R. (1980). The dynamics of composing: Making plans and juggling constraints. In L. Gregg and E. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 31-50). Hillsdale, NJ: Erlbaum.

Fuller, M. (1975, June). Teaching the process of problem solving. Paper presented at Annual Conference, American Society for Engineering Education, Colorado State University, Fort Collins, Colorado.

Fuller, M. (1978). Reasoning for experimenters. Paper presented at Frontiers in Education Conference, Orlando, Florida.

Gelman, R., and Gallistel, C.R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Ginsburg, H.A. (1977). *Children's arithmetic: The learning process*. New York: Van Nostrand Reinhold.

Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist, 39*, 93-104.

Groen, G., and Parkman, J.M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79*, 329-343.

Haynes, J.R. (1981). *The complete problem solver*. Philadelphia: Franklin Institute Press.

Hetherington, E.M. (1983). Socialization, personality, and social development. (Vol. IV of P.H. Mussen, Ed., *Handbook of child psychology*. New York: Wiley.

Hiebert, J., and Wearne, D. (1985). A model of students' decimal computation procedures. *Cognition and Instruction, 2*, 175-205.

Hursh, B., Haas, P., and Moore, M. (1983). An interdisciplinary model to implement general education. *Journal of Higher Education, 54* (1), 42-59.

Johnson, R.H. (1981). The new logic course: The state of the art in non-formal methods of argument analysis. *Teaching Philosophy, 4* (2), 123-143.

Johnson, R.H., and Blair, J.A. (1980). The recent development of informal logic. In J.A. Blair and R.H. Johnson (Eds.), *Informal logic* (pp. 3-28). Pt. Reyes, CA: Edgepress.

Jones, B.F., Amiran, M., and Katims, M. (1985). Teaching cognitive strategies and text structures within language arts programs. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 259-296). Hillsdale, NJ: Erlbaum.

Jones, B.F., Friedman, L.B., Tinzmann, M., and Cox, B.E. (1984). *Content-driven comprehension instruction and assessment: A model for Army-training literature*. (Tech. Rep.). Alexandria, VA: Army Research Institute.

Journal of Educational Psychology. (1921). Intelligence and its measurement: A symposium. *The Journal, 12*, 123-147, 195-216.

Just, M.A., and Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354.

Kintsch, W. (1979). On modeling comprehension. *Educational Psychologist, 14*, 3-14.

Kruglanski, A.W. (1981). Endogenous attribution and intrinsic motivation. In M.R. Lepper and D. Greene (Eds.), *The hidden costs of reward* (pp. 85-108). Hillsdale, NJ: Erlbaum.

Larkin, J.H., McDermott, J., Simon, D.P., and Simon, H. (1980). Expert and novice performance in solving physics problems. *Science, 208*, 1335-1342.

Lazere, D. (1982). *Composition for critical thinking: A course description*. San Luis Obispo, CA: California Polytechnical State University Press.

Lepper, M.R. (1981). Intrinsic and extrinsic motivation in children: Detrimental effects of superfluous social controls. In W.A. Collins (Ed.), *Minnesota symposium on child psychology* (Vol. 14, pp. 155-214). Hillsdale, NJ: Erlbaum.

Lepper, M.R. (1983). Extrinsic reward and intrinsic motivation: Implications for the classroom. In J.M. Levine and M.C. Wang (Eds.), *Teacher and student perceptions: Implications for learning* (pp. 281-318). Hillsdale, NJ: Erlbaum.

Lepper, M.R., and Greene, D. (1981). Overjustification research and beyond: Toward a means-end analysis of intrinsic and extrinsic motivation. In M.R. Lepper and D. Greene (Eds.), *The hidden costs of reward* (pp. 109-148). Hillsdale, NJ: Erlbaum.

Lipman, M. (1982). *Harry Stottlemeier's discovery*. Philadelphia: Temple University Press. (Original work published in 1974).

Lipman, M. (1985). Thinking skills fostered by philosophy for children. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 83-108). Hillsdale, NJ: Erlbaum.

Lochhead, J. (1985). Teaching analytical reasoning skills through pair problem solving. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 109-132). Hillsdale, NJ: Erlbaum.

Mandl, H., Stein, N.L., and Trabasso, T. (Eds.). (1984). *Learning and comprehension of text*. Hillsdale, NJ: Erlbaum.

Matz, M. (1982). Towards a process model for high school algebra errors. In D. Sleeman and J.S. Brown (Eds.), *Intelligent tutoring systems* (pp. 25-50). New York: Academic Press.

McGraw, K.O. (1981). The detrimental effects of reward on performance: A literature review and a prediction model. In M.R. Lepper and D. Greene (Eds.), *The hidden costs of reward* (pp. 33-60). Hillsdale, NJ: Erlbaum.

McPeck, J. (1981). *Critical thinking and education*. Oxford, England: Martin Robinson.

Meichenbaum, D. (1985). Teaching thinking: A cognitive behavioral perspective. In S.F. Chipman, J.W. Segal, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 2. Research and open questions* (pp. 407-426). Hillsdale, NJ: Erlbaum.

Messick, S. (Ed.). (1976). *Individuality in learning*. San Francisco: Jossey-Bass.

National Assessment of Educational Progress (NAEP). (1983). *The third national mathematics assessment: Results, trends and issues* (13-MA-01). Denver, CO: Educational Commission of the States.

Neves, D.M., and Anderson, J.R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

Newell, A., and Estes, W.K. (1983). Cognitive science and artificial intelligence. In Committee on Science, Engineering, and Public Policy (Ed.), *Research briefings 1983*. Washington, DC: National Academy Press.

Newell, A., and Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-56). Hillsdale, NJ: Erlbaum.

Nicholls, J.G. (1983). Conceptions of ability and achievement motivation: A theory and its implications for education. In S.G. Paris, G.M. Olson, and H.W. Stevenson (Eds.), *Learning and motivation in the classroom* (pp. 211-238). Hillsdale, NJ: Erlbaum.

Nickerson, R.S., Perkins, D., Smith, E.E. (1985). *The teaching of thinking*. Hillsdale, NJ: Erlbaum.

Norris, S.P. (1985). Thinking about critical thinking: Philosophers can't go it alone. In D.A. Nyberg (Ed.), *Philosophy of education*.

Palincsar, A.S., and Brown, A.L. (1984). Reciprocal teaching and comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*, 117-175.

Papert, S. (1980). *Mindstorms*. New York: Basic Books.

Paris, S.G., and Jacobs, J.E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development, 55,* 2083-2093.

Paris, S.G., Cross, D.R., and Lipson, M.Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76,* 1239-1252.

Paul, R. (1982). Teaching critical thinking in the "strong" sense: A focus on self-deception, world views, and a dialectical mode of analysis. *Informal Logic, 4,* 3-7.

Paul, R. (In press). A review of Critical thinking and education. *Informal Logic.*

Perfetti, C. (1985). *Reading ability.* New York: Oxford University Press.

Perkins, D.N. (1982). Difficulties in everyday reasoning and their change with education. *Final Report to the Spencer Foundation, Project Zero,* Harvard University.

Peterson, P., Swing, S., Stark, K., and Waas, G. (1984). Student's cognitions and time on task during mathematics instruction. *American Educational Research Journal, 21,* 487-515.

Polson, P.G., and Jeffries, R. (1985). Analysis—instruction in general problem-solving skills: An analysis of four approaches. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 417-455). Hillsdale, NJ: Erlbaum.

Reif, F., and St. John, M. (1979). Teaching physicists' thinking skills in the laboratory. *American Journal of Physics, 47,* 950-957.

Resnick, D.P. (1980). Minimum competency testing historically considered. *Review of Research in Education, 8,* 3-29.

Resnick, D.P., and Resnick, L.B. (1977). The nature of literacy: An historical exploration. *Harvard Educational Review, 47,* 370-385.

Resnick, L.B. (1987). The development of mathematical intuition. In M. Perlmutter (Ed.), *Minnesota symposium on child psychology* (Vol. 19, pp. 159-194). Hillsdale, NJ: Erlbaum.

Resnick, L.B., Cauzinille-Marmeche, E., and Mathieu, J. (1987). Understanding algebra. In J. Sloboda and D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 169-203). New York: Oxford University Press.

Rubenstein, M.F. (1980). A decade of experience in teaching and interdisciplinary problem-solving course. In D.T. Tuma and F. Reif (Eds.), *Problem solving and education: Issues in teaching and research* (pp. 25-38). Hillsdale, NJ: Erlbaum.

Scardamalia, M., and Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S.F. Chipman, J.W. Segal, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 2. Research and open questions* (pp. 563-578). Hillsdale, NJ: Erlbaum.

Schoenfeld, A. (1982). Measures of problem-solving performance and of problem-solving instruction. *Journal for Research in Mathematics Education, 13,* 31-49.

Schoenfeld, A. (1983). Problem solving in the mathematics curriculum: A report, recommendations, and an annotated bibliography. *Report prepared for The Mathematical Association of America Committee on the Teaching of Undergraduate Mathematics.*

Schoenfeld, A. (1985). *Mathematical problem solving.* New York: Academic Press.

Segal, J.W., Chipman, S.F., and Glaser, R. (Eds.). (1985). *Thinking and learning skills: Vol. 1. Relating instruction to research.* Hillsdale, NJ: Erlbaum.

Simon, H.A. (1976). Identifying basic abilities underlying intelligent performance of complex tasks. In L.B. Resnick (Ed.), *The nature of intelligence* (pp. 65-98). Hillsdale, NJ: Erlbaum.

Sleeman, D. (1983). Assessing aspects of competence in basic algebra. In D. Sleeman and J.S. Brown (Eds.), *Intelligent tutoring systems* (pp. 185-199). New York: Academic Press.

Stein, N.L. (1986). Knowledge and process in the acquisition of writing skills. In E.Z. Rothkopf (Ed.), *Review of research in education, 13* (pp. 255-258). Washington, DC: American Educational Research Association.

Stein, N.L., and Trabasso, T. (1982). What's in a story: An approach to comprehension and instruction. In R. Glaser (Ed.), *Advance in instructional psychology* (Vol. 2, pp. 213-267). Hillsdale, NJ: Erlbaum.

Sternberg, R.J. (1986). *Intelligence applied.* New York: Harcourt Brace Jovanovich.

Sternberg, R.J., and Detterman, D.K. (Eds.). (1979). *Human intelligence: Perspectives on its theory and measurement.* Norwood, NJ: Ablex.

Svenson, O., and Hedenborg, M. (1979). Strategies used by children when solving simple subtractions. *Acta Psychologica, 43,* 477-489.

Thorndike, E.L., and Woodworth, R.S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 3,* 247-261, 553-564.

Toulmin, S.E., Reike, R., and Janik, A. (1979). *An introduction to reasoning.* New York: Macmillan.

Tuma, D.T., and Reif, F. (Eds.). (1980). *Problem solving and education: Issues in teaching and research.* Hillsdale, NJ: Erlbaum.

Turkewitz, H., O'Leary, K.D., and Ironsmith, M. (1975). Generalization and maintenance of appropriate behavior through self-control. *Journal of Consulting and Clinical Psychology, 43* (4), 577-583.

vanDijk, T.A., and Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Voss, J.F., Greene, T.R., Post, T.A., and Penner, B.C. (1983). Problem-solving in the social sciences. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research theory* (Vol. 17, pp. 165-213). New York: Academic Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, Eds. and Trans.). Cambridge, MA: Harvard University Press.

Wales, C.E. (1979). Does how you teach make a difference? *Engineering Education, 69,* 394-398.

Wales, C.E., and Nardi, A.H. (1985). Teaching decision-making: What to teach and how to teach it. In A.L. Costa (Ed.), *Developing minds: A resource book for teaching thinking.* Alexandria, VA: Association for Supervision and Curriculum Development.

Wales, C.E., and Stager, R.A. (1977). *Guided design.* Morgantown, WV: West Virginia University Center for Guided Design.

Weinstein, C.E., and Underwood, V.L. (1985). Learning strategies: The how of learning. In J.W. Segal, S.F. Chipman, and R. Glaser (Eds.), *Thinking and learning skills: Vol. 1. Relating instruction to research* (pp. 241-258). Hillsdale, NJ: Erlbaum.

Whimbey, A., and Lochhead, J. (1982). *Problem solving and comprehension.* Philadelphia: The Franklin Institute Press.

Whimbey, A., and Lochhead, J. (1984). *Beyond problem solving and comprehension.* Philadelphia: The Franklin Institute Press.

Wickelgren, W.A. (1974). *How to solve problems.* San Francisco: Freeman and Co.

Woods, D.R. (1983). Introducing explicit training in problem solving into our courses. *Higher Education Research and Development, 2,* 79-102.

Woods, D.R., Crowe, C.M., Taylor, P.A., and Wood, P.E. (1984). The MPS program for explicitly developing problem solving skill. In L.P. Grayson and J.M. Biedenbach (Eds.), *ASEE Annual Conference Proceedings* (Vol. 3, pp. 1021-1035). Salt Lake City, UT: American Society for Engineering Education.

Woods, S.S., Resnick, L.B., and Groen, G.J. (1975). An experimental test of five process models for subtraction. *Journal of Educational Psychology, 67,* 17-21.

SCIENCE EDUCATION: IT'S NOT ENOUGH TO 'DO' OR 'RELATE'

Kathleen J. Roth

Elementary science teaching has not changed much in the last forty years; it was, and is, basically fact oriented and didactic. Textbooks present science as a large body of knowledge and science learning involves memorizing facts about dinosaurs, planets, weather, etc. A crowded elementary curriculum and elementary teachers' lack of expertise in science pose an additional, enduring problem: Very little instructional time is allotted to science teaching. Reformers have not been satisfied with a science curriculum that emphasizes recall of facts, and this dissatisfaction has led to a series of attempts to change the nature of elementary science teaching. Each of these efforts has attempted to help students develop higher-order thinking skills—to think scientifically.

By far the most serious and best supported of these reform attempts was the inquiry movement in the 1960s and 1970s, which gave rise to extensive National Science Foundation programs in curriculum development and teacher education. By the 1980s, however, it was apparent that the inquiry movement had failed to achieve its goals. Teaching practice and commercial materials had adopted some of the rhetoric of inquiry teaching but had not changed their basic character and had not had a significant impact on student learning outcomes.

Thus in the 1980s the reform movement split into three groups: (a) a still-powerful group that continues to advocate inquiry teaching, (b) a Science-Technology-Society (STS) group that focuses on changing the goals and content of science teaching, and (c) a conceptual change group that focuses on changing the methods of instruction so that they are more responsive to students' thinking and development. All of these groups agree on the need for reform and on goals that emphasize the development of conceptual understanding and higher-level thinking skills. Advocates of the three perspectives, however, differ in their definition of the relationships among content, process, and attitude goals; in their analyses of the nature of scientific thinking; and in their recommendations about what elementary children can and should be taught about science and the nature of scientific thinking.

Each of the three perspectives contrasts with traditional elementary science teaching in that they each select particular goals for focus and suggest a framework to help teachers

---

Reprinted with permission from the Winter 1989 issue of The American Educator, the quarterly journal of the American Federation of Teachers.

focus on meaningful outcomes, instead of trying to "cover it all." While it is tempting to think about ways to blend the three perspectives, in this analysis I will stress the *distinctions* among them. I argue that elementary teachers need a framework or instructional model that helps them limit and focus their teaching on meaningful, achievable learner outcomes; telling teachers to "do it all" is counterproductive. An eclectic approach to science instruction (throw in a little bit of everything) is all too evident, I fear, in most elementary science textbooks in use today.

A brief description of typical textbook-based instruction will be used as a point of contrast with the three alternative perspectives. To provide additional points of contrast, examples of how each perspective might organize instruction to teach fifth graders about plants and photosynthesis will be described.

Traditional Textbook-Based Science Teaching

The traditional method of teaching science is focused on textbook instruction and didactic teaching. This content-mastery approach to science is organized around discrete topics—planets, electricity, magnetism, dinosaurs—with little attempt to make connections across topics. The instructional pattern typically consists of reading the text followed by answering factual questions posed by either the teacher or the text. Hands-on activities and teacher demonstrations are added to foster motivation but are often selected because they are easy to do or fun rather than for their usefulness in developing conceptual understanding or higher-level thinking. Thus, science teaching (and scientific thinking) is viewed as student acquisition of facts.

In the Silver Burdett & Ginn *Science* series (Mallinson, Mallinson, Valentino, & Smallwood, 1989), for example, fifth graders learn about photosynthesis in a chapter that presents a smorgasbord of information about activities of green plants. In five lessons, students learn about life processes; transportation of materials in plants; functions of roots, stems, and leaves; structure of leaf and stem cells (veins, stomata, chloroplasts, chlorophyll); the process of photosynthesis (role of energy, water, carbon dioxide, products); storage of manufactured food in fruits and vegetables; use of food energy in respiration; comparison of respiration and photosynthesis; and the use of energy for reproduction (parts of the flower, pollination, fertilization, germination). Thus, a wide spectrum of content is covered at a rapid pace.

In the lesson about photosynthesis, students read detailed information about the cell structure of leaves. The five steps in the photosynthetic process are listed and summarized in a word equation (water + carbon dioxide + energy —> sugar + oxygen). To "evaluate student understanding," teachers are directed to ask factual questions that can be answered in short phrases taken directly from the text: What does a plant need to make food? (water, carbon dioxide, light energy) (Mallinson et al., 1989, p. 12).

In a suggested enrichment activity, students look at leaf cells under a microscope. They draw cells and label the chloroplasts. In this lesson on photosynthesis, students read a lot of information and reproduce it in small bits when prompted by teacher questions. Only in a review question at the end are students required to explain anything (How do green plants make food?). The lesson does not ask students to use photosynthesis to explain everyday observations of plants or even to explain their observations in the enrichment activities.

The lesson on photosynthesis is followed by a lesson on plant respiration and a lesson on plant reproduction. This sequence of lessons does not seem likely to foster conceptual understanding and higher-level thinking. Instead, instruction is a parade of lessons marching along at a brisk pace, inundating observers (students) with a panoply of facts and concepts to be absorbed.

Inquiry Perspective

The inquiry perspective contends that students will develop better understandings of the nature of science and will be more interested in science if they are engaged in "doing" science. Student investigations of phenomena (not textbooks) are the backbone of the curriculum, and the focus of these investigations is on the use and development of science inquiry or process skills—predicting, hypothesizing, observing, recording data, making inferences and generalizations, etc. Students are viewed as little scientists who explore phenomena through hands-on activities and who use and develop scientific thinking skills to build up knowledge and conceptual understandings in the same ways that scientists use experimental work to construct new knowledge, concepts, and theories.

The proponents of an inquiry orientation acknowledge that science content and science thinking processes are both important, interrelated parts of science. They place a clear emphasis on the science thinking skills, however, allowing science content to play a secondary role. In the Elementary Science Study (ESS) materials (Educational Development Center, 1971), content is seen as almost irrelevant. The important thing is that children are engaged in scientific thinking and actions: asking questions, manipulating materials, observing phenomena, and making up explanations to answer their questions. Students "mess about" with mealworms or batteries or mystery powders, and the particular facts or concepts that they learn from these activities are not nearly as important as the lessons they learn about the nature of scientific inquiry.

In an inquiry perspective, the science process skills are generally regarded as the heart of scientific thinking, and these skills are typically organized in a hierarchical fashion. For example, the developers of *Science—A Process Approach* (SAPA) identified thirteen science processes and organized them in a hierarchy according to students' cognitive development (Commission on Science Education of the AAAS, 1968). Processes such as observing and classifying were seen as appropriate to teach at the early elementary level. Integrated process skills (hypothesizing, controlling variables, interpreting data, etc.) were deemed more complex and appropriate for the intermediate grades. Others organize these skills in a hierarchy that matches inductive models of a scientific method.

What is striking in the inquiry-oriented literature on higher-level thinking skills in science is the careful attention paid to describing and sequencing the science thinking process skills and the relative isolation of these thinking skills from a conceptual context. Little attention is given to the nature and development of scientific conceptual understanding. There is an underlying assumption that scientists' development of conceptual knowledge is a straightforward process: that conceptual knowledge is simply the "product" of the scientific thinking processes.

In an inquiry-oriented curriculum, science process skills are the main goal of elementary science teaching. Higher-level thinking is promoted by engaging students in using science processes in hands-on investigations. Content knowledge, conceptual understanding, and positive attitudes toward science are viewed as outgrowths of the inquiry investigations. The focus of instruction is on the development of process skills. If students can understand and use these general thinking process skills, they will be able to develop meaningful conceptual understanding in any area of study.

Photosynthesis is the focus of the fifth-grade "Producers" unit in the SCIIS *Communities* Teacher's Guide (Knott, Lawson, Karplus, Thier, & Montgomery, 1978). In this unit, three major activities serve as the focus of classroom lessons, and no textbook is used. First, students germinate and measure the growth of various seed parts. This activity is designed to illustrate that germinating seed embryos get food from the seed.

In a second activity, students plant grass seeds and keep some in the light and some in the dark to demonstrate that plants need light to grow and to suggest that plants do not get food from the soil. Toward the end of this activity, the teacher explains photosynthesis to the students and the experiment is interpreted in light of this explanation. Finally, students germinate and measure the growth of bean plants under various conditions: with and without cotyledons (the developing embryo's food supply in the seed) and with and without light. Students are expected to use the idea of photosynthesis to explain their results.

Ms. Kain (pseudonym) and three other teachers whom we observed in our study (Roth, 1984; Smith & Anderson, 1984) spent six to eight weeks teaching this unit. In each of these classrooms, the bulk of instructional time was spent on setting up the experiments, measuring plants, recording results on a class scatter plot, and using the scatter plots to average data and draw line graphs to show patterns of growth. Discussions were also important aspects of lessons. Ms. Kain's questions focused on eliciting students' predictions, observations, and explanations of the experiments. She rarely gave out information, explaining photosynthesis only once during the unit and once in a unit review. Thus, Ms. Kain created many opportunities and a safe environment for students to construct their own explanations of observed phenomena; she never drilled students about definitions or details of the photosynthetic process. She spent eight weeks exploring photosynthesis with her students (approximately twenty-four lessons), quite a contrast with the one-lesson coverage of the Silver Burdett textbook.

And what did the students learn? Our post-tests (Roth, Smith & Anderson, 1983) assessed students' conceptual understanding by asking a variety of questions about how plants get their food. Eighty-nine percent of the students in the study failed to grasp the central concept of the unit: that plants get their energy-containing food only by making it internally out of carbon dioxide and water. Instead, most students began and ended the unit believing that plants take in food from the outside environment and that plants, like people, have many different kinds of food (air, water, fertilizer, minerals, soil, sun, etc.). Students watched and measured plants growing in the light and the dark, they conducted the experiments with germinating seedlings, but they interpreted these observations in terms of their preconceived ideas and failed to integrate the teacher's presentations of photosynthesis into their interpretations of the experiments. At best, students added plants' making of food as one of *many* sources of food for plants, failing to understand the *unique* ability of green plants to use light energy to convert non-energy-containing raw materials into energy-containing food that is necessary to support growth and life functions. A similar pattern was seen in textbook-focused classrooms. Thus, an inquiry approach did not provide any advantage in promoting conceptual understanding.

And what did students learn about the nature of scientific inquiry, scientific processes, and scientific attitudes? Many of them found "doing" science (the science processes) "fun" but ended up frustrated by the focus on processes and by all the measuring and recording of data. As Rachel explained:

> I don't know *why* we kept measuring those plants. I mean it was fun for awhile, but I already know that plants need light, and now I know it again.

What did Rachel learn about science processes and scientific thinking? She learned that it involves a lot of activity that does not help you make any better sense of things. She learned that science activities and processes are ends in themselves. It is important, for example, to make careful observations and to record them accurately not because such care helps you develop better understanding but because "that's what you do in science." Because Rachel did not develop better conceptual understanding the processes of science seemed meaningless and not worth the effort. Driver (1983) critiques this doing of science in the absence of meaningful conceptual development, suggesting that the "I do and I understand" slogan might more appropriately be "I do and I am even more confused."

The results of this study also suggest that involvement in hands-on activities did not produce the desired student *attitudes* toward science. Although the activities made science seem fun, they did not necessarily help students value and feel comfortable with science. As a learner, Rachel was clearly frustrated that the doing of science did not lead her to any better personal understandings about plants' need for light. Although she was in a classroom environment where her ideas were valued and where she felt comfortable sharing her ideas, Rachel did not leave the unit feeling good about herself as a learner of science or comfortable in the neighborhood of science. She wondered why she held the same understandings at the end of the unit that she had held at the beginning. She had spent eight weeks measuring, observing, and talking about plants, and no change in her understanding had occurred! This was not a satisfying learning experience for her nor was it an experience that will make her enthusiastic about studying more science. These findings suggest that assessments of student attitudes toward science must dig beneath the surface; it is not enough to know that students have positive attitudes toward hands-on activities. What are they learning about science and scientific thinking?

Science-Technology-Society Perspective

Many science educators advocate a dramatic change in the goals of K-12 science education. They argue that the overarching purpose of school science is not to create future scientists but to create citizens who understand science in multidimensional, multidisciplinary ways that will enable them to participate intelligently in critical thinking, problem solving, and decision making about how science and technology are used to change society (environmental issues, nuclear power, personal health, energy resources, etc.). Yager and Hofstein (1986) describe this perspective in sharp contrast to both the traditional, disciplinary-based focus of science teaching and to the inquiry approach:

> In some respects the traditional content and process dimensions of science may be the dimensions least important and appropriate to us in planning for the year

> 2000. They may be least important for helping us attain a scientifically and technologically literate citizenry for which so many yearn. If so, they may be the dimensions of science that deserve little or no emphasis as a science curriculum is planned and newly conceived for all K-12 students.

Yager and Hofstein suggest six essential goals of a quality K-12 science curriculum. These goals stand in clear contrast with those espoused by the inquiry perspective:

> 1. The human being, human potential, human advances, and human adaptions will serve as the organizer of the curriculum (instead of the structure of the disciplines or the scientific processes).
>
> 2. Current problems and societal issues will serve as the backbone of the curriculum.
>
> 3. Science and technological processes that students can use in everyday life will be emphasized over processes that scientists use.
>
> 4. Practice with decision-making skills using science and technology knowledge in a relevant, social context will be emphasized over skills needed to "uncover correct answers to discipline-bound problems."
>
> 5. Awareness should be an integral part of science learning.
>
> 6. In dealing with problems and issues, ethical, moral, and value dimensions will be considered (in contrast with traditional science instruction, which is taught as value free and discipline bound).

Thus, an STS curriculum is human and society focused, problem centered, and responsive to local issues. Problems to be investigated are selected for their relevance to students' lives and their multidisciplinary nature. As in the inquiry perspective, students are seen as active learners, but the activities they engage in are focused on *using* scientific and technological knowledge to solve problems and make decisions rather than on *creating* scientific knowledge. Thus, in the STS perspective, students act as young science citizens rather than young research scientists.

The STS perspective shares with many of the inquiry programs a tendency to separate process goals from content goals and to emphasize process goals over content goals. Content is to be selected based on its interest and relevance for students and on the richness of the societal problem it provides. Thus, content is selected for its potential to serve the primary goals of developing students' decision-making and problem-solving process skills and of helping students learn to integrate values and moral thinking in this decision-making process.

Thus, process or thinking skills are of primary importance in the STS perspective, but the process skills are of primary importance in the STS perspective, but the process skills are defined quite differently from those in inquiry programs. The processes emphasized in the STS perspective focus on wise *use* of scientific knowledge in decision making and problem solving about societal problems rather than on the construction of scientific knowledge through careful observation, inferencing, experimentation, etc.

In the STS perspective, decision making and problem solving are viewed as the higher-level thinking skills that students should develop. These skills are seen as needed by all students, not just those bound for science careers. To "think scientifically," then, one does not need to be able to create or discover scientific ideas the way a research scientist does. Rather, one needs to be a good consumer of scientific knowledge—finding and using scientific ideas as needed to solve particular problems.

As in the inquiry perspective, there is a striking lack of analysis in the STS literature about the particular content or concepts needed to make good decisions or to solve problems. The kinds of conceptual knowledge needed to do effective decision making and problem solving have not been carefully analyzed or defined. This lack of attention to the nature of conceptual understanding suggests an assumption that the conceptual knowledge needed to think scientifically (to make decisions and to solve problems) is straightforward information that students will rather easily incorporate into their own thinking. Thus, the challenge (or higher-level aspects) of scientific thinking is not in understanding concepts but in using concepts to make decisions and solve problems.

Since materials from an STS perspective are only now in the development phase, I will speculate about a possible STS unit on photosynthesis. A science-technology-society unit would probably not teach about photosynthesis as the focus of a unit. Instead, photosynthesis would be addressed in the context of exploring a technological or scientific problem facing society. For example, a unit might be structured around the problem of the effects of deforestation and industrialization on the warming of the Earth's atmosphere due to the greenhouse effect. Science concepts relevant to this problem include absorption of solar energy by the Earth's atmosphere and the changing balance of O^2-CO^2 in the Earth's atmosphere due to fossil fuel use and widespread cutting of rain forests (plants use carbon dioxide in the photosynthetic process and release oxygen). The teacher would help students use these concepts to assess the severity of the problem and to think about possible solutions to the growing danger of the greenhouse effect: Should the chopping down of rain forests be slowed down? How? Why would developing nations resist pressures to slow deforestation? What are other ways of slowing the greenhouse effect? How can local citizens influence decisions about slowing fossil fuel usage?

Unit activities might focus on role playing and assessment of the problems from different points of view. Scientific processes could be investigated in the context of studying the evidence that the greenhouse effect is actually a danger. Students might read arguments from scientists holding different opinions about the severity of the threat. The unit-culminating activity would be a student-generated activity designed to take action on these issues, such as writing and circulating a pamphlet about ways to reduce energy consumption in the home, studying home gas and electric bills and trying to decrease consumption for a month, writing to state representatives or members of Congress in support of particular bills related to energy issues, and planting trees on school grounds and encouraging others to plant trees.

In this approach, students would be engaged in complex thinking that requires them first to understand and evaluate the soundness of the evidence offered by expert scientists and then to integrate those understandings with understandings of social and political processes. For the unit to be meaningful and to result in student generation of a worthwhile project, the role of the teacher would be a very complex one. Teachers would have to understand underlying scientific principles, the evidence supporting scientists' predictions of the greenhouse effect, and the relationships between political and scientific issues.

The knowledge that students are likely to learn about photosynthesis in such a unit would be limited in the traditional disciplinary sense. For example, discussions of photosynthesis in such a unit would be limited in the traditional disciplinary sense. For example, discussions of photosynthesis are likely to focus on CO^2-O^2 balance rather than on the food-making function. Because the unit activities focus on problem solving and citizenship action, students are more likely to end the unit understanding that extensive cutting of forests and burning of fossil fuels may result in warming climates worldwide rather than remembering much about photosynthesis. Helping students see how science and technology can both cause problems and help solve problems would be another challenge in teaching this unit. Without such an appreciation, students might come to view science and technology as evil and threatening. Thus, they would not value scientific inquiry and knowledge generated by science.

Conceptual Change Perspective

Of the three perspectives, the conceptual change perspective is the only one that did not originate in response to a social/political climate that demanded reform in science education. In contrast, it is a research-based perspective that grew out of cognitive science studies of learning and knowing in knowledge-rich domains. Research on teaching and learning in science, in particular, provided critical insights about why it is so difficult for students to develop useful, conceptual understandings of many of the subjects they are taught in school (Anderson and Roth, in press; West and Pines, 1985).

In a conceptual change perspective, the primary goal of science education is to help students develop meaningful, conceptual understandings of science and its ways of describing, predicting, explaining, and controlling natural phenomena (Anderson and Roth, in press; Driver, 1987). In this view, scientific knowledge is meaningful to learners only when it is *useful* in making sense of the world they encounter.

Scientific knowledge that can be used by learners is characterized by rich connections between concepts and facts and is organized around key ideas in ways that make the knowledge accessible and able to provide broad explanatory power. This stands in contrast to knowledge that exists as isolated fragments that students can parrot back for recall-focused tests but cannot apply in explaining real-world phenomena. Such connected knowledge is not locked into one tightly organized structure that simply gets larger as a learner adds new knowledge into it (a passive, additive view of learning and a static view of knowledge). Rather, this set of connected knowledge is flexible and constantly changing as the learner revises, reorganizes, and deepens understandings over time (an active, conceptual change view of learning and knowledge growth).

This web of knowledge, or the individual's conceptual ecology (Posner et al., 1982), only becomes useful and meaningful to students when it is integrated with the learner's own personal knowledge and experiences with natural phenomena. Students come to science classes with many ideas and explanations about natural phenomena. Their ideas are experience based and often stand in stark contrast to the scientific explanations studied in school. A central goal of science teaching is to help students *change* their intuitive, everyday ways of explaining the world around them—to incorporate scientific concepts and ways of thinking into their personal frameworks.

Like the STS advocates, advocates of a conceptual change perspective argue that a central goal of science education is to develop scientific literacy for *all* students. The two perspectives, however, define the nature of a scientifically literate society in different ways and they suggest different instructional approaches. The conceptual change perspective focuses on the power of conceptual understanding, arguing that the way to develop scientifically literate citizens is to change instruction so that it helps learners develop rich and meaningful understandings of whatever science they study. Such understanding will help students value science as a sense-making endeavor. The STS perspective, in contrast, approaches the need for a scientific citizenry from a curricular perspective, arguing that it is changes in the *content* of the curriculum that are critically needed. To capture the interests of all students, traditional disciplinary-bound curricula need to be abandoned in favor of current science and society issues.

Unlike the inquiry and STS perspectives, the conceptual change perspective views conceptual knowledge as central in science and in science learning. Where do the scientific processes fit in this scheme? Scientists' work, this position argues, is conceptually driven, and the so-called scientific processes cannot be separated from scientific conceptualizations. Science teaching, therefore, should integrate science processes and conceptual knowledge in ways that better reflect the richness and complexity of science itself.

What does it mean to say that scientists' work is "conceptually driven"? Expert scientists do not hypothesize, make inferences, or design experiments in the absence of conceptual frameworks. Their conceptual frameworks are not only influenced by their observations and inferences; their frameworks also *drive* and shape the hypotheses they make, the questions they raise, the things they pay attention to in their observations. What distinguishes their work as science is not these processes, which are processes that are equally applicable in history, economics, mathematics, or the arts, but the particular knowledge that organizes how these processes are used. A scientist who observes well, for example, is not one who spends endless hours documenting and describing every possible detail that can be observed about a particular phenomenon. Instead, a good scientific observation focuses on key features in ways that will contribute new knowledge, that increases the explanatory power of a particular conceptual framework, that generates new understandings about relationships among concepts, or that raises significant questions about accepted conceptual frameworks. The scientist who makes such critical observations does not just happen to see what others have not seen because of expert "observation skills." To set up the conditions in which important observations are made, the scientist draws from existing conceptual knowledge, asks questions about important pieces of the framework, develops hypotheses, and designs experiments that will permit the critical and relevant observations to be made. Thus, skillful observation is intimately linked with both conceptual knowledge and with all the other "processes" of science. As Millar and Driver (1987) argue, each of the scientific "processes" can be submitted to a similar critique. These processes are not meaningful in isolation, and they are not science in the absence of a scientific conceptual framework.

Thus, the integration of science processes and conceptual knowledge is characteristic of expert scientists' thinking. In addition, this approach maps more effectively onto students' ways of learning science. When science processes are taught in relative isolation from conceptual knowledge, students often develop a misleading picture of science as "doing" for doing's sake. In my own studies of process-focused teaching, I identified students whose naive theories went unchallenged and unchanged despite weeks of process-focused instruction. These students did not view the processes of science as helping them better understand their world. Scientific processes, and science itself, were viewed as mysterious things that only certain smart people do and understand.

Practicing process skills in isolation of conceptual development does not help students understand how science processes are useful in understanding the world. Instruction that involves students in using these scientific processes to change their own theories in ways that are personally meaningful and consistent with scientific explanations provides a powerful alternative to process-focused instruction. To make such changes, students need to do difficult cognitive work that includes the use of science processes (predicting, hypothesizing, observing, inferring). Students do not, however, practice these processes in isolation and the goal is not for them to be better observers or predictors. Instead, these processes are used in the service of developing better explanations of natural phenomena.

The conceptual change perspective challenges two traditional ways of characterizing "higher-level thinking" in science. First, as discussed above, the usual representation of conceptual knowledge and science processes as separate kinds of scientific thinking does not accurately reflect what we have learned about the nature of science learning. Conceptual knowledge and the scientific processes must be intimately interwoven to be meaningful. Second, views of scientific thinking processes as a hierarchy of levels suggest that conceptual understanding is a relatively straightforward, lower-level process that students need as a basis for doing higher-level thinking. This view does not capture the ways in which the processes typically labelled as "higher order" (application, synthesis) are used in the development of meaningful, conceptual understandings.

In my own studies of middle school learners who successfully integrated ideas about photosynthesis into their personal conceptions about how plants get their food (Roth, 1986), I gained a new appreciation of the challenges students face in changing their personal theories and misconceptions and coming to understand scientific concepts. The successful conceptual change learners used a variety of "higher-level" thinking skills in the process of developing the conceptual understanding that plants make (rather than take in) their own food. These students continually tried to use new information to explain and make predictions about plants and their needs. Often this was a difficult and confusing process, as students encountered areas where their knowledge was incomplete or in conflict with ideas presented ("I thought *water* was the plants' food"). Such conflicts led to analyses of the differences between them and to restructuring of students' personal conceptual frameworks. Students used experimental observations as well as teacher explanations to help them rethink their ideas. These students were metacognitively active, monitoring their developing understandings and resolving areas of confusion ("If water isn't their food, then why does my Mom water her plants?"). In the end, they developed conceptual understandings of photosynthesis that they could use to explain why plants in a cave will die or why we need to water plants and give them sunlight. Their ability to apply these concepts to a variety of real-world phenomena, however, was a *reflection* of their conceptual understanding rather than a higher-level thinking skill that was developed after their conceptual understanding was in place (as suggested by taxonomies like Bloom's that place "application" at a higher level than "knowledge"). These students' conceptual learning was not straightforward, bounded, lower-level thinking. Their conceptual change learning required the integration of a variety of sophisticated, higher-level cognitive and metacognitive thinking skills. Thus, the development of meaningful conceptual knowledge is a complex process, involving a great deal of higher-order thinking. In a conceptual change perspective, conceptual understanding *is* higher-level thinking.

Recently, I taught a group of fifth-grade students about photosynthesis using a modified version of conceptual change curriculum material I had written and evaluated in two earlier research studies (Roth, 1985; Roth & Anderson, 1987). The materials began by presenting a problem to be solved: What is food for plants? Only after students' entering answers to this question had been elicited and challenged was the idea of photosynthesis explained. Students were then give numerous opportunities to make sense—to reconcile the idea of photosynthesis with their entering conceptions. Photosynthesis was taught as a first step in helping students learn in a year-long program about ecological interactions and about changes in energy and matter.

The unit began by eliciting students' definitions of "food." The text explained the difference between everyday definitions of food (anything we eat or drink) and a definition of food as energy-containing matter. After agreeing that water is not food by this definition, the students became involved in a lively debate about how plants get their food. Most students asserted that water is one important food for plants, while others pointed out that water does not have energy in it. Students had interesting experience-based arguments to support their belief that water did provide energy for *plants* ("rain water has energy in it but drinking water doesn't"). The class made a chart listing their ideas about how plants get their food, and this chart was revisited several times throughout the month-long unit.

Two hands-on experiments (adapted from the SCIIS "Producers" unit, Knott et al., 1978) and one discussion of a historical experiment provided initial information to help students puzzle through the problem of how plants get their food. An experiment with germinating seed parts suggested that young plant embryos get food that is stored in the seed's cotyledon. In discussing the results of this experiment, the idea that water did not contain energy for plants was again discussed. But many students still made predictions in the next experiment that grass seeds would grow in both light and dark conditions if they were watered because water is their food. Water was still a confusing issue.

While the grass seeds were given time to grow, von Helmont's famous experiment of 1642 was read and discussed. Students' predictions agreed with von Helmont's: As the tree in the tub of soil grows, the weight of the soil will decrease as the tree "eats" materials from the soil. When the students found out that the soil in von Helmont's experiment did not lose any significant weight, they accepted this as evidence that the plant was not "eating" soil. But they remained puzzled about why plants need soil and what fertilizers and minerals do for plants. I encouraged these questions while trying to keep the students focused on the problem of how plants get energy-containing food. As the grass plant experiment proceeded, and plants in the dark began to yellow and die, students became increasingly convinced that sunlight was somehow important and unsure about the role of water. Clearly, soil and water *alone* could not keep a plant alive. The sun seemed to be very important. But is the sun itself the food for plants, as many students believed?

Only when most of the students' entering ideas about food for plants had been either ruled out or brought into question was an explanation of photosynthesis given: that cells in plants' leaves use light energy from the sun to change water and air into energy-containing food. After this idea was explained in contrast with students' entering ideas, students were given many opportunities to use photosynthesis to explain everyday phenomena. In addition, their own application questions were rewarded and encouraged. One student, for example, wondered about the fact that *only* plants could make food. Did that mean that without plants, there would be no food? Another countered that we could live on candy bars if we didn't have plants. I encouraged this debate and later brought in Snickers bars. We analyzed the list of ingredients, talking about how each ingredient (corn syrup, sugar, peanuts, chocolate) had been made by a particular kind of plant.

Application opportunities included problems posed in overhead transparencies (How did this plant produce a potato?), questions posed for students to write and/or talk about (What if only one leaf of a plant could get sunlight? Could it live? Why or why not?), and further experimentation. Students also wrote in science log books about their evolving ideas, receiving feedback questions and comments from the teacher. Application activities included analysis of controlled experiments as well as other, seemingly nonscientific activities, such as role playing the life of a bean seed embryo. Card-sorting activities where students constructed (in pairs and individually) different concept maps showing the relationships among ideas were also used.

At the end of the unit, students revisited their initial explanations of how plants got their food and wrote and talked about how their ideas had changed. The post-test asked a series of application questions for which students had to write out predictions and explanations. There was also a mini-interview question in which students were asked to arrange words related to photosynthesis (air, water, sun, fertilizers, soil, cotyledon, leaf, stem, etc.) in a conceptual map and then explain their arrangement. Most students demonstrated a coherent understanding of the key concepts; the interviewer questions permitted individual coaching to clarify remaining confusions. Students' comments and questions revealed ways in which their conceptual understanding had generated a new sense of wonder and sense making:

> **John:** Ms. Roth, I used to think that plants were just kind of *there*, ya' know? They just sat there. But now I know that they're really very busy little things, aren't they? There's lots going on inside them.
>
> **Ted:** I know plants can use water, air, and sun to make food...but, I mean, *how* do they do that?

This unit involved much more time than the one lesson on photosynthesis outlined in the Silver Burdett & Ginn text (Mallinson et al., 1989). And yet the unit did not present many of the technical terms covered in the Silver Burdett & Ginn lesson—chloroplasts, stomata, carbon dioxide, hydrogen. But students were able to make predictions and observations, to change and develop explanations, and to apply ideas in a meaningful time frame; that is, they were provided time to change their understandings. The instructional model proved both workable for me and productive of meaningful conceptual understanding for my students. Although it did not call for explicit teaching of science processes, students regularly used these processes in constructing their developing understandings. Although it did not focus on exploration of societal issues, the unit provided a base for later study of the cutting down of rain forests explored during a unit on ecosystems. And students came to appreciate an issue important in their own society—that we need plants to produce *all* of our food (including Snickers!).

Conclusions

Each of the perspectives reviewed in this paper uses analyses of the nature of scientific thinking to define the goals of elementary science instruction and to advocate particular approaches to science teaching. The three perspectives share a concern that science teaching practice needs to change: It needs to become more focused on helping students think scientifically rather than on memorizing facts. To accomplish this goal, difficult curricular and instructional decisions have to be made. Given elementary teachers' limited time for planning and teaching of science, it is critical that the scope of the elementary science curriculum be limited and that teachers be supported in developing a coherent and manageable model or framework to guide their science planning and teaching. I find the conceptual change perspective to provide the most promise in terms of impact on students' understanding of science and scientific thinking. This framework also provides a useful tool for guiding daily teaching and planning because it frees teachers from long lists and hierarchies of skills and facts to be learned and helps them focus instead on important aspects of scientific understanding.

References

Anderson, C.W., & Roth, K.J. (In press). Teaching for Meaningful and Self-Regulated Learning of Science. In J. Brophy (Ed.), *Advances in Research on Teaching* (Volume 1). Greenwich, CT: JAI Press.

Black, P.J. (1987). The school science curriculum: Principles for a framework. In A.B. Champagne and L.E. Hornig (Eds.), *The science curriculum: The report of the 1986 National Forum for School Science.* Wash., DC: AAAS.

Commission on Science Education of the American Association for the Advancement of Science (AAAS) (1986). *AAAS science—A process approach.* Stamford, CT: Xerox.

Driver, R. (1987). Promoting conceptual change in classroom settings: The experience of the Children's Learning in Science Project. In J.D. Novak (Ed.), *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics.* Ithaca, NY: Cornell University.

Educational Development Center (1971). *The elementary science study.* Manchester, MO: Webster Division, McGraw-Hill Book Co.

Knott, R., Lawson, C., Karplus, R., Their, H., & Montgomery, M. (1978). *SCIIS communities teacher's guide.* Chicago: Rand McNally.

Mallinson, G.G., Mallinson, J.B., Valentino, C., & Smallwood, W.I. (1989). *Science,* Morristown, NJ: Silver Burdett & Ginn.

Millar, R., & Driver, R. (1987). *Beyond processes.* Studies in science education, 14, 33-62.

Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science education, 66*(2), 211-227.

Roth, K.J. (1985). *Food for plants teacher's guide* (Research Series No. 153). East Lansing, MI: Institute for Research on Teaching, Michigan State University.

Roth, K.J. (1986). *Conceptual-change learning and student processing of science texts* (Research Series No. 167). East Lansing, MI: Institute for Research on Teaching, Michigan State University.

Roth, K.J., & Anderson, C.W. (1987). *The power plant: Teacher's guide to photosynthesis*

(Occasional Paper No. 112). East Lansing, MI: Institute for Research on Teaching, Michigan State University.

Roth, K.J., Smith, E.L., & Anderson, C.W. (1983, April). *Students' conceptions of photosynthesis and food for plants*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Smith, E.L., & Anderson, C.W. (1984). Plants as producers: A case study of elementary science teaching. *Journal of Research in Science Teaching, 21*, 685-698.

West, L.H.T. & Pines, A.L. (1985). *Cognitive structure and conceptual change*. New York: Academic Press, Inc.

Yager, R.E., & Hofstein, A. (1986). *Features of a quality curriculum for school science*. Journal of Curriculum Studies, 18(2), 133-146.

LEARNING IS CONNECTING NEW INFORMATION TO PRIOR KNOWLEDGE

Research from numerous sources---philosophy of science, psychology, education---all seem to indicate that learners do not come to learning situations as "blank slates," but rather learners possess knowledge that is used, reorganized, modified, or discarded in the learning process. The six articles in this section draw from the diverse research that has examined the "act" of learning.

The first two articles, written by philosophers von Glasersfeld and Matthews respectively, present contrasting views on "constructivism." Simply put, constructivist argue that students build up---construct---their own knowledge. Students are engaged actively in the learning process. Also, learning takes place within a social and cultural environment and is, therefore, socially mediated. Von Glasersfeld states that knowledge can be constructed by combining and recombining learned concepts and trying out new relationships among them. In his article, von Glasersfeld responds to common questions about constructivism and discusses educational implications of adopting a radical constructivist view. A key observation he makes concerning science education is that hypotheses, natural laws, and other scientific theories should not be taught as "truth" but rather as the "most viable" explanations for phenomena. Matthews' article presents a complex philosophical argument against constructivist perspectives such as von Glasersfeld's. Matthews primarily argues that constructivism fails to account for the real objects of science (or, more accurately, of the world), such as falling apples, which are to be contrasted with the theoretical objects, such as gravity, a system of mathematically expressed relations. As he states, "constructivism is correct in stressing the inventive, humane, culturally and temporally dependent aspects of creating theoretical aspects of science---but none of this bears, of itself, upon truth."

The article by Songer and Linn presents the results of research on how students integrate knowledge. They studied 153 eighth-graders and identified their beliefs about the nature of science. Then Songer and Linn determined whether a relationship existed between students' beliefs and their ability to integrate knowledge. Finally, Songer and Linn had the students try principles and prototypes as loci of knowledge integration and indicate whether these loci were useful to them and which they preferred. The authors found that students with a "dynamic" view of science acquire a more integrated understanding of scientific principles than those who possess "static" views. Songer and Linn's article primarily provides evidence for students' active construction of knowledge and warns against "focusing science instruction too narrowly on facts or isolated pieces of scientific knowledge."

The last three articles fall under the perspective of "conceptual change" (introduced in the previous article by Roth). The articles are premised on the notion that students come to learning situations with misconceptions (or as Aldridge calls them, preconceptions). The article by Posner , Strike, Hewson, and Gertzog presents the principles of conceptual change. For students to change their current set of conceptions about natural phenomena, they must develop dissatisfaction with that set of conceptions; they must confront or uncover an alternative conception that seems intelligible and plausible; and then, gradually, they must make a commitment to the alternative conception as they find it fruitful in various applications. The articles by Anderson and Minstrell offer very specific evidence and advice on how to bring about conceptual change in the classroom. Anderson suggests that science concepts can be put in the context of meaningful problems. Students can then ask their own questions, realize their own misconceptions, and gradually reconstruct their understanding of how the world works. Minstrell presents a sample physics lesson using conceptual change principles---most notably relying on the strategy of conceptual conflict. He then discusses more generally what the conceptual change principles suggest for helping students "restructure" knowledge, for the design of instruction, and for fostering an appropriate classroom environment for learning.

QUESTIONS AND ANSWERS ABOUT RADICAL CONSTRUCTIVISM

Ernst von Glasersfeld

Introductory Remark

The questions I try to answer in the pages that follow are those that I was asked at a NARST meeting (Atlanta, April 1990) and at a AAAS symposium (Washington, DC, February 1991). Reviewing the questions later, I found that they could be roughly divided into three subject areas. I begin with the specifically epistemological questions, go on to those concerned with the problem of social interaction, and end with some implications the constructivist orientation might have for teachers and the philosophy of instruction. Since the answers I give are not derived from an established dogma but spring from my subjective point of view, the reader will find a certain amount of overlap between the three sections. I like to claim that this is inevitable because, in my experience, once one shifts to the constructivist orientation, everything one thinks and does changes in a way that seems remarkably similar and coherent. But let me emphasize a point I have made in many of my papers:

Constructivism, as far as I am concerned, is one possible way of thinking. It is a model, and models, no matter how useful they might prove, must never be claimed to be "true."

Epistemology

Q1. ***Is constructivism primarily an epistemology or a pedagogy?***

A1. Constructivism deals with questions of knowledge—what knowledge is and where it comes from. It can therefore be considered an exercise in epistemology. The idea that knowledge is constructed by the knower is as old as Western philosophy. The pre-Socratics in the 6th and 5th century B.C. were already aware that not everything they knew could be said to come from sensation and that the mind, the specifically human mind, was to a large extent responsible for shaping human knowledge. Some of them also noticed that there was a serious problem with the claim that knowledge could and should faithfully reflect a world held to be independent of the knower. They had

---

From "Questions and Answers About Radical Constructivism," by Ernst von Glasersfeld, in ***The Practice of Constructivism in Science and Mathematics Education****, Kenneth Tobin, Editor. Washington, DC: AAAS Press, in preparation. Used with permission.*

realized there is no way of checking knowledge against what it was supposed to represent. One can compare knowledge only with other knowledge. The later skeptics have never tired of reiterating this irrefutable argument.

Recently, the reviewer of a paper of mine said that constructivism was "post-epistemological," and she was quite right (Nel Noddings in "Constructivist views on the teaching and learning of mathematics," JRME Monograph #4, edited by Bob Davis, Carolyn Maher, and Nel Noddings, NCTM, 1990). Indeed, the constructivist theory does not fit into the conceptual patterns of traditional epistemology, precisely because it posits a different relation between knowledge and that "real" outside world. Because it changes that important relation, I prefer to call it a theory of knowing rather than a theory of knowledge. (Please note that constructivism does not deny an outside world; it merely agrees with the skeptics and holds that the only world we can know is the world of our experience.)

Q2. ***What's the difference between constructivism and radical constructivism? How do you define radical, trivial, and social constructivism?***

A2. A few years ago, when the term "constructivism" began to become fashionable and was adopted by people who had no intention of changing their epistemological orientation, I introduced the term "trivial constructivism" to distinguish this fashion from the "radical" movement that broke with the tradition of cognitive representationism.

"Social constructivism" is a recent development, started by some who claim radical constructivism does not take into account the role of social interaction in the construction of knowledge. The claim, as I see it, is partly justified by the fact that neither Piaget nor any more recent constructivist has actually specified a detailed model of how social interaction works from the constructivism point of view. On the other hand, Piaget and all of us who have derived our impetus from him, have always maintained that social interaction is a powerful influence in the construction of knowledge. But we were busy devising models for all the elementary constructing that has to be done before a cognitive organism can begin to know and interact with others. As far as I know, the social constructivists tend to take society as a given, and a radical constructivist cannot accept this. From my point of view, "society" must be analyzed as a conceptual construct before its role in the further construction of concepts can be explained and properly assessed.

Q3. ***What are the differences between constructivism and idealism? Is the distinction a qualitative one?***

A3. Idealism holds that the mind and its constructs are the only reality. Constructivism holds that we can know only what our minds construct, but that its constructing is not free (ad lib): just as some ways of acting "work" and others do not, so some of our conceptual constructs and theories work and others do not. For constructivism, the fact that something works (is "viable") does not mean that it therefore is a representation of that "real" world that prevents other things from working. The "real" world remains unknowable no matter how well we manage in the domain of our experience. (This, by the way, also differentiates constructivism from pragmatism with which it has much in common.)

Q4. ***Was Descartes revealing his constructivist view when he said "I think therefore I am."?***

A4. Descartes was a very complex thinker and not averse to the notion of construction. Analytic geometry, one of his major achievements, is an explicit example of cognitive construction. But I would not interpret his famous "cogito" phrase as a constructivist statement about the self. The mind was very much an ontological given for Descartes. He called it <I>res cogitans<$> (the thinking thing), and he saw it as an entity separate from a world that exists in its own right (<I>res extensa<$>) and about which the knower can find out some things, not through the senses but by logical reasoning.

For constructivists the "I," the agent that does the constructing, is a construct like everything we claim to know. This makes the model circular and upsets traditional philosophers but not cyberneticians.

Q5. ***What is the role of truth in your model?***

A5. Constructivism is an attempt to cut loose from the philosophical tradition and specifically from the philosophical tradition that knowledge has to be a representation of reality, where reality is spelled with a capital and means a world prior to having been experienced. Truth in constructivism, as I keep repeating, is replaced by viability. That is a stark step to take. I've given many philosophical arguments (or excuses, if you will) for taking this step in a number of papers. Here let me just mention one thing, because it seems the most important in this context. If we want to talk in terms of "Truth" and mean by this that what we say and what we think should be a replica of the world as it is, then we have to believe that we can visualize or imagine what that "real" world is like. We have to conceive of that world in terms of existence in the sense that it exists by itself prior to our noticing, perceiving, and thinking about it. Now, I claim that we cannot put any meaning to the expression "to exist" outside our sphere of experience. To us, "to exist" means by and large to have a location in space and in time. But if space and time, as Kant said, are forms of our experiencing and do not actually belong to the ontological reality beyond our experience, we have no way of visualizing things without space and time. And for me that is a good reason for giving up the notion that we have to know or imagine what the world is like before we experience it. Rather we must try to develop a theory that provides a relatively coherent explanation for how we come to have the kind of knowledge we use in our living.

Q6. ***Is knowledge something to be presented to a learner or is it an activity?***

A6. Knowledge is always the result of a constructive activity and, therefore, it cannot be transferred to a passive receiver. It has to be actively built up by every single knower. A teacher, however, can orient a learner in a general direction, and constraints can be set up that prevent the learner from constructing in directions that seem unsuitable to the teacher.

Q7. ***Doesn't the fact that we can predict physical phenomena with a great deal of accuracy mean that the picture of reality we have constructed is congruent with the "real" world outside?***

A7. No. This question goes to a key element of constructivism. Predictions we make, whether we are physicists or ordinary human beings, predict experiences, that is, things we perceive, feel, and think. What we perceive, feel, and think is necessarily the result of our ways and means of carrying out these activities. The results can never be said to be like, let alone "congruent," with an ontological reality; all we may conclude is that this "real world" allows us to perceive and think certain things. If a prediction turns out to be right, a constructivist can only say that the knowledge from which the prediction was derived proved viable under the particular circumstances of the case.

Q8. ***Some concepts, "chair" for example, seem to be based on a physical stimulus, while others, such as "I.Q.," seem more abstract. Is there a continuum from those triggered by some external stimulus to those that are more abstract?***

A8. When we say "physical stimulus," we say more than we know. All we actually know is that we have had, or are having, a perceptual or, as Piaget would say, a sensory-motor experience (see A7). To conclude that, because we have a perceptual experience which we call "chair," there must be a chair in the "real" world, is to commit the realist fallacy. We have no way of knowing what is or could be beyond our experiential interface. If we can reliably repeat the chair experience, we can only conclude that, under the particular circumstances, it is a viable construct.

Something like "I.Q." is not a sensory-motor experience but, like any measurement, the result of a number of abstractions. In Piaget's terms, it would be an "operative" concept, because it is derived from operations we have to carry out. The terms "concrete" and "abstract" are not useful in constructivism, because conceptual analysis shows that there are many levels of abstraction. The first level would be abstraction from sensory-motor experiences. Once such an abstraction has been made, its result immediately can be taken as material for a further abstraction, and so on. Thus, there is a continuum of levels of abstraction.

Science consists to a large extent of relational (operative) concepts that are the result of various abstractions that have their origin in sensory-motor experience and our own mental operations. This has been beautifully expressed by Einstein in his essay on "Physics and reality" (in *Ideas and opinions*, Bonanza Books, 1954). There, he explicitly says that our object concepts are "free creations of the human (or animal) mind."

Q9. ***Are there physical scientists who disagree with constructivism? If so, name one or two and indicate where they disagree with what you have said.***

A9. Among the great physicists, most have said and written things that imply a constructivist position. Nearly all these statements were anticipated by Hermann von Helmholtz, who wrote in 1881: "The principle of causality is in fact nothing but the presupposition of the lawlikeness of all of the appearances of nature" (von Helmholtz, Epistemological writings, Dordrecht: Reidel, 1977; p. 180). I feel, however, that contemporary physicists are in tacit agreement with Popper, who admits that the skeptics' arguments are irrefutable and that an "instrumentalist" view of knowledge is quite reasonable, but goes on to say that it is better for scientists to continue to believe that they are getting closer to understanding the "real" world. Needless to say, I do not share this double standard.

From the limited experience I have had in the last few years with physics instruction, I would say that most physics teachers have little sympathy for constructivism. They prefer to talk as though they were describing an absolute reality. If pushed a little, they may say, "Well, of course, we don't know for sure," or "Students don't like uncertainty, they want to know how things really are."

Q10. ***Is there a concept of progress in radical constructivism, or of no change, or is all change random?***

A10. Constructivists would be rather foolish to deny that we know more today than did, say, the pre-Socratics or even Newton. But the growth of knowledge concerns knowledge of how to do things. The fact that we can "send a man to the moon," or we can have a phone conversation from our moving car, does not mean that the conceptual structures or action programs we have developed are any more a representation of ontological reality. We have learned a great deal about how to set up and manage useful regularities in our experiential world. We know many more things the world allows us to do. We have, after all, learned even how to demolish our experiential world irrevocably. But for constructivism it is an illusion to believe that what we can do gives a picture of the "real" world. It describes our experiential reality as we happen to be experiencing it. And what we experience is necessarily shaped and coordinated according to the concepts and conceptual relations we are using at the moment.

Q11. ***Is truth then an act of faith?***

A11. Yes. Constructivism does not deal with the traditional conception of "Truth," which would require that one knows, or at least believes, that an idea, a theory, or any conceptual construct is an accurate representation or duplication of something beyond our experiential field. Truth in that sense may be felt (perhaps acquired) by following the mystics, but it is not a rational procedure that leads to it. Note that I am in no way disparaging or discounting what the mystics tell us. I am merely saying don't try to explain it rationally.

Constructivism replaces the notion of truth with that of viability, which does not refer to anything outside the experiential field. This is analogous to pragmatism; but pragmatists formulated the slogan "true is what works," and their use of the word "true" is, to my mind, misleading.

Q12. ***If we cannot know reality, can we understand, test, and evaluate others' constructed realities? How can we test another's scientific hypothesis and have valid grounds for rejecting or accepting it?***

A12. Scientific hypotheses are "tested" in the experiential world where they either do or do not what they are claimed to do. Experiential worlds belong to individuals, but in the course of social interaction these individual worlds become adapted to one another and come to form a consensual domain, an area where the interactors' mutual expectations are more or less regularly realized (see "The Role of the Social," below). The consensual domain of the scientific subculture, because of its many specific constraints, is particularly homogeneous, and new hypotheses are "tested" against that (relatively) homogeneous background. If a new hypothesis makes too many new assumptions, it will usually be discarded regardless of whether it works or not (for example, Wegener's hypothesis of "continental drift," proposed at a time when the notion of floating continents was considered preposterous by the entire scientific community. It took a lot of new experiential "evidence" to become acceptable.)

Q13. ***Does reality exist aside from one's construction of reality? How do you know?***

A13. I don't. As far as logic is concerned, I could be one of Leibniz' monads, that is, a closed, impermeable entity whose "experience" is like a movie that automatically runs its preordained course. But this is not a particularly useful model. I prefer the model built on the notion of viability within the space left free between constraints whose origin I can never tell.

The main difficulty of the question arises from the word "exist." In our human usage, it means to have some location in space, or time, or both. But since space and time are our experiential constructs, "to exist" has no meaning outside the field of our experience, and whatever an independent ontological reality may do, it is not something we can visualize or understand.

Q14. ***Shouldn't there be a new label for reality as it exists, to distinguish it from subjective reality?***

A14. Yes. But a label is not much use if one can never demonstrate what it stands for. I try to say "experiential reality" as often as I can and, in contrast, put that other mystical "reality" in quotation marks (see Q11). In German there are two convenient words, "Wirklichkeit," the reality due to one's actions, and "Realitat" which can be used for the other.

Q15. ***What harm is there in remaining a realist? What do I have to gain by thinking as a radical constructivist?***

A15. No harm at all, as long as you don't tell others that the reality you have constructed is the one they ought to, or worse, must believe in. In everyday interactions, we all go on as though our experiential reality were the same obligatory reality for everyone. But if one gets into psychology or education, taking account of the subjective nature of experiential realities becomes a matter of honesty.

Q16. ***How does constructivism escape the "hermeneutic circle": "Is constructivism an adequate view of knowledge and the world despite the fact that all knowledge is held to be relative?"***

A16. Hermeneutics searches for the "true" interpretation, constructivism does not. Constructivism is deliberately circular and never claims "objectivity." We cannot look at our experiential worlds from the outside. We construct them from the inside and have usually lived a good many years in them, before we begin to wonder where they came from and what they "really" are. It would be more than a little presumptuous to believe that we could "know" the world that has produced us and our ways and means of thinking. It would require what Hilary Putnam has called "the God's eye view." In contrast, we can always decide whether a way of acting or thinking is "adequate," because it either does what we expect of it, or it does not.

Q17. ***Is constructivism a belief?***

A17. If you call it "belief" when one has constructed a model and finds it the most adequate one has so far developed, then I would say yes. But it is certainly not belief in an absolute "truth" or revelation. To me, it simply seems to be the most viable view at the moment.

Q18. ***You alluded to the existence of an all knowing "real" God. How can you reconcile the belief in constructivism with the belief in a real God?***

A18. "Alluding to God" is not the same as believing in one. Vico, the first explicit constructivist, said "A human being can know only what humans have made, God can know the world because He created it." This is not even an hypothesis but merely a conjecture. It might be the case and it might not be the case. A mystic may "know," but that knowledge would not be rational knowledge. This is not a statement of value but a statement of methodology. I have the greatest respect for mystics although I know nothing about how they operate.

Q19. ***How do you hold up constructivism for scrutiny?***

A19. By letting people see whether it works, where "to work" means to be as consistent as possible and to cover as much ground as possible. The assessment has to be in terms of viability not in terms of "truth."

Q20. ***Is realism typically a Western philosophy? If so, would you say something about Eastern philosophy.***

A20. Our textbooks and television (even NOVA programs) are all too often manifestations of scientism, a pseudo-philosophy that I would call the most recent Western religion precisely because it claims to have found objective "truth." The realism the scientistic religion entails is certainly characteristic of the West.

The little I have learned about Eastern philosophy does not make me competent to answer the second part of the question. I have the impression that Buddhism and Taoism tend to take rational knowledge and commonsense reality less seriously than we do, and the teaching of Zen seems to aim precisely at weaning the student from the frantic attachment to naive realism.

The Role of the Social

Q21. ***How does constructivism view a body of common knowledge, such as in a textbook or field?***

A21. It is an illusion that there is knowledge in textbooks or documents. They contain language—strings of words—put into them by authors. The words have meaning for both the authors and the readers and interpreters, each one of whom has built up his or her subjective meanings according to his or her individual experience. Though these individual meanings are constructs that have been through a certain amount of social adaptation (because their users have socially interacted with others), they remain subjective and to some extent idiosyncratic, which never fails to come to the surface in a serious discussion. Texts contain neither meaning or knowledge; they are scaffolding on which readers can build their interpretation (see Q23, below).

Q22. ***Where and how is language taught?***

A22. I would say that the foundation of language is not taught at all. Children learn that by themselves. You can't teach language to a two year old or a three year old. You may occasionally be able to show them the use of a new word, but even this often does not work too well. Later, specific interaction can focus on specific linguistic difficulties. That is why teaching language at a later stage by correcting what children say is enormously important. Unfortunately this seems to have gone by the board, because parents no longer have many extensive conversations with their children. Conversation has been replaced by television, and television, being a moneymaking enterprise keyed to the lowest levels of literacy, rarely provides a model for good language.

But then, you might say, a lot of language can be learned from dictionaries. Dictionaries, however, contain definitions that are simply other words, they do not provide the experience from which meanings are abstracted. They do not give us new cognitive elements, but only new combinations of elements that we already possess.

In general, language is learned in the course of interaction with other speakers, because speaking is a form of interacting, and we modify our use of words and utterances when they do not yield the expected results. Insofar as a classroom is interactive it contributes to the students' linguistic development and also provides them with opportunities to witness the use of words in the context of the experiences to which they refer. That is why it is so very important to provide to students the experiences about which they are to learn.

Q23. ***What is the role of "social instruction"?***

A23. I take it this refers to "group learning," and there is a lot to be said about this, including the following points:

(1) Students who work at a problem together with other students have to verbalize how they see the problem and what they want to do about it. This is one way of generating reflection, which requires awareness of what one is thinking and doing. This, in turn, provides occasions for active abstraction (repeating, writing down, and learning by heart what a teacher says, does none of this).

(2) Explaining something to a peer usually leads to seeing things more clearly and often to spotting inconsistencies in one's own thoughts. And when a small group explains its "solution" (irrespective of whether it happens to be viable or not) to the whole class, this provides a wonderful opportunity for learning (compare the work of Paul Cobb, Ken Tobin, Grayson Wheatley, Jack Lochhead, and others).

(3) Knowing that those you work with have no ready made answer increases everyone's courage to try and find one.

(4) If one of the group finds an answer, this more often than not generates motivation to try a new problem.

(5) To have an inconsistency or "error" explained by a peer is far less painful than have the teacher tell you that you are wrong.

Q24. ***Please give examples how social interaction would promote understanding.***

A24. Once again my examples will have to be simplistic. The first kind of example would be the failure to get what one asks for or what one is trying to get or to get the wrong thing. I'm talking about very young children and their acquisition of linguistic knowledge. Anyone who has attentively observed a child's language acquisition knows the situation where the child uses a word because it wants something, but the word is inappropriate and therefore does not do what the child expects. The adult misunderstands, gives the wrong thing, and the child gets very irritated. Such moments of failure and perturbation may lead to a change. The child may try another word or the parent may suggest one. Then the child may try the new word the next time and if it works, the child will have learned. It's a very simple inductive procedure. You hang on to the schemes, the gambits, and the methods that have worked in the past. We all do this, not only in the area of language but in all areas of learning. And this principle has an important corollary that I have often mentioned: We have no reason whatsoever to change what we are doing as long as it produces the desired result. This is an assertion that links with the notice of misconceptions in physics or in other teachings. If a conception has satisfied what the students demanded of it up to now, the students have no earthly reason to change it. The love of truth will not make them change it. You have to show them that their conception has limitations and that there are situations where that conception doesn't work. That is one way of getting them interested in learning a new conception, because they all want to widen the range of application of whatever schemes or ideas they have.

Q25. ***How do we reach shared meanings?***

A25. The expression "shared meaning" is misleading. Your meaning and another's are at best compatible. Which is to say, in a given situation, neither reacts in a way that the other could not expect. One only has to get into a discussion of philosophy to realize just how tenuous this compatibility is on the level of abstract concepts. It usually takes a long evening and a good deal of patience before one gets even a dim idea of what the other is trying to say. In a research team, for example, it can take well over a year to establish a workable compatibility with regard to the main terms that are being used. This is what we have called the generation of consensual domain.

Q26. ***How can we teach the tentativeness and personal nature of knowledge, yet not undermine the cultural and social reality we have arrived at?***

A26. We cannot. But, as a constructivist, I feel it is one's duty to proclaim the "tentativeness" of everything one teaches and to warn the students that most, if not all, of the "facts" one presents will be seen in a different light two or three decades later. What matters is the process of construction and the fact that this process yields results that seem to be viable at the moment. The cultural and social reality would be a far more livable and fruitful one if we could do away with the notion that we have The Truth and others had better believe it.

Q27. ***What is the point of talking or lecturing?***

A27. Well, in a way, I think, I have answered this. The point is to foster new combinations of concepts. By talking to an audience, I cannot give people any new concepts, but I can prod them to combine in different ways the concepts that they have associated with the words I am using. I can propose combinations they may not have thought of before or have had no occasion to use before. And I can perhaps generate some perturbation about conceptual structures that have been used unthinkingly merely because they felt familiar. If I am not causing some conceptual change, my talking or writing is useless. But in saying this, I am already in the topic of teaching.

Instruction

Q28. ***What are the implications of constructivism for a theory of instruction?***

A28. There are many. Here I can mention only a few that can be stated briefly.

1) If we assume that students have to build up their own knowledge, we have to consider that they are not "blank slates." Even first graders have lived for a few years and found many viable ways of dealing with their experiential environment. The knowledge they have is the only basis on which they can build more. Hence it is crucial for the teacher to get some idea of where they are (what concepts they seem to have and how they relate them).

2) Whatever a student says in answer to a question (or "problem") is what makes sense to the student at that moment. It has to be taken seriously as such, regardless of how odd or "wrong".it might seem to the teacher. For the student, to be told that it is wrong, is most discouraging and inhibiting. Besides, given the way the student interpreted the question, the answer may actually be a good one.

3) If a teacher wants to modify a student's concepts and conceptual structures, he or she has to try and build up a model of the particular student's own thinking. Models of students' thinking can of course be generalized, but before assuming that a student fits the general pattern one should have some solid evidence that this is a viable assumption in the particular case. It should never be assumed that students' ways of thinking are simple or transparent.

4) Asking students how they got to the answer they gave is a good way of discovering something about their thinking and opens the way to explaining why a particular answer may not be useful under different circumstances.

5) If you want to foster students' motivation to delve further into questions that at first are of no particular interest from the students' point of view, you will have to create situations where the students have an opportunity to experience the pleasure inherent in solving a problem. Simply being told an answer is "correct" does not do anything for the student's conceptual development if he or she was not interested in the question.

6) Successful thinking is far more important than "correct" answers. Successful thinking should be rewarded even if it was based on unacceptable premises.

7) To understand and appreciate students' thinking, the teacher must have an almost infinitely flexible mind (because students sometimes start from premises that seem inconceivable to teachers).

8) Constructivist teachers can never justify what they teach by claiming that it is "true." In mathematics they can show that it derives from certain conventional operations; in science, they can not say more than that it is the best way of conceiving the situation because, at the moment, it's the most effective way of dealing with it.

Q29. ***How do you activate the mind to construct knowledge?***

A29. You activate students' minds to construct knowledge by letting them struggle with problems of their own choice, helping them (as Maria Montessori said a long time ago) only when they ask for help. At best, the teacher can orient a students' constructing in a fruitful direction, she or he can never force it. This is, of course, time consuming; but after students have experienced once or twice the pleasure of finding a solution by their own thinking, they will be quite ready to work on problems suggested by the teacher.

Q30. ***What is the role of training interactions? Can episodes of training be used to solicit episodes of teaching?***

A30. Making the distinction between training and teaching does not mean that there is no place for training. I am convinced that some training is very useful, but in itself it is not conducive to understanding. However, there are many experiences for which one needs to be proficient in certain routines. Routines are necessary. You are not going to be a mathematician unless you have automated some of the elementary operations; but if you have automated them without having grasped their conceptual underpinnings, you are not going to be a mathematician either. Training and teaching have to go hand in hand. If I mostly emphasize teaching, it is because training is common practice in schools anyway.

Q31. ***How does constructivism influence problem solving?***

A31. By letting the students discover that it can be fun to solve problems. It will not be fun if the teacher constantly watches that they take the "right" path to the "right" solution. Students often take wholly unexpected, unconventional paths to a solution that is quite viable but perhaps limited in its applicability. If the teachers don't respect this, they merely demonstrate that they are unaware of the history that has led to the present state of the art. That is why, especially in science, I consider it indispensable that students be given a feeling for the long history of magnificent failures.

Q32. ***Is knowledge really only built from existing pieces? And does meaning only come from prior knowledge?***

A32. I think I have already answered this, at least in my way. Knowledge is often built up by combining and recombining available concepts or by trying out new conceptual relations. After the child has acquired a certain level of language, language can be used to suggest and to stimulate new combinations, and so on. And a similar succession of levels can be seen in reflection and abstraction, the processes that generate mathematics and other abstract realms.

Q33. ***Are the implications of constructivism the same for mathematics and science education?***

A33. Most are, but some are not (see A28, #8). In mathematics, for instance, the teacher should explain that the step in calculus that leads from infinitesimal bits to continuity is a conceptual decision, not a logical consequence; just as the decimal system with all its consequences and implications is a conceptual system, not a God given characteristic of an external "reality." In science, on the other hand, the first thing to make clear is that scientists do not "unveil" the objective order of a preexisting universe, but invent viable ways of coordinating and managing experiences, where the range of experiences is always limited by the particular interests of the given period. The history of science is full of examples that show how theories are superseded when new experiences enter the field of exploration. Old theories, as a rule, are not proven to be "wrong"; they merely turn out to be inadequate in an expanded domain of experience or in the pursuit of new goals. Above all, the constructivist teacher should never present a solution as the only one.

Q34. ***How can we structure a teacher education program that allows students to construct an understanding of science, teaching, and learners that is consistent with constructivism?***

A34. By making clear that education pursues two goals. First, to foster thinking that does not involve conceptual contradictions and leads to internally consistent results. Second, to introduce students to the consensual domain that governs the discipline at the moment. The "science" of today is, after all, what the scientists of today believe they agree on. But to a constructivist, "agreement" does not entail that the concepts and conceptual relations used by the people who "agree" are identical; it merely means that, under the given circumstances, they seem compatible (which, to repeat myself, indicates no more than that one is unable to discover discrepancies in one's own and the other's way of speaking and acting).

Q35. ***If we adopt the radical constructivism view of knowledge, what changes in the way science teachers talk about the world will be required (that is, in language or terminology)?***

A35. Little beyond the fact that what is being talked about is a particular view of the world, not the world as such. It is a way of ordering and coordinating experience that seems useful and consistent at the moment.

Q36. ***Technical terms are intensively employed in the teaching and learning of science; what is the status of technical terms (which are rigidly defined) in scientific understanding? Technical terms are memorized, right?***

A36. That is a very good question. Technical terms are somewhat different from the ordinary words that children learn in the first three or four years of their lives. The reason they are different is that there usually is a good operational definition for technical terms, and the student can be shown or guided to do what is necessary to construct the referent. The operations to produce a viable referent may be physical or they may be conceptual, and in the second case, teaching the term will be more difficult. The physicists' notion of "energy" is a good example. For the last two years I have been involved with people who do research on physics teaching. As a nonphysicist I was at first astonished to see that physics teachers and physicists quite often begin to argue about what energy actually is and that there does not seem to be a definition that satisfies everyone. Of course, there are definitions in the textbooks, but the physicists tend to say that these definitions reflect a very naive way of thinking about energy. So there are, indeed, difficulties, and a nonphysicist is not in a position to solve them. All I would say is this: It is no use presenting students with a verbal definition unless they have an opportunity to have some kind of relevant experience. In a laboratory one can show all sorts of simple experiments in which a construct such as energy becomes, if not obvious, at least quite plausible. So much for technical terms.

Q37. ***Does constructivism favor divergent instructional strategies? Or is it consistent with a teacher having an agenda?***

A37. The "agenda," I would say, is the attainment of the twofold goal (see Q34 above): instructional strategies may be suggested but not prescribed. Every inspired teacher will develop his or her own. Teaching, after all, is something of an art. It requires not only acquaintance with techniques but also intuition.

Q38. ***If understanding is in the individual, does the teacher then abdicate any evaluative function?***

A38. My answer is no. In fact, the teacher cannot help evaluating what students do and say. Teachers must show the student what they consider inadequate, and they must show why it does not work. The constructivist teacher is not concerned with "truth" but with successful operating. It is no use simply telling students that they are wrong. This only decimates such motivation as they may have. To change their way of acting or thinking, they have to see in their own experience that what they did was not the most successful way of operating.

Q39. ***Whose construction is to be accepted in a school setting?***

A39. The first criterion to decide the acceptability of a "construction" is its viability, that is, whether it does or does not do what it is supposed to do. If it is viable, one can introduce other criteria, such as simplicity, economy, elegance, convention, and so on.

Q40. ***What are the implications of constructivism for how we assess student learning?***

A40. This is the most difficult question. Before it could be answered satisfactorily, one would have to decide what one wants to teach. As long as evaluation is based on "performance," where performance means the ability to reproduce answers that were part of the instructional process, we are not assessing what is the important "learning" from a constructivist point of view, that is, the ability to solve new problems. We have to decide whether we want students to develop their ability to think or simply to handle standard procedures and to give standard answers in a limited set of situations. Given sufficient authority, the second can be achieved by rote learning and does not involve students' conceptual ability. If they pass this kind of test, they may find the satisfaction of holding a job, and they will have to discover other ways of attaining intellectual satisfaction. Much research is going on in the area of "assessment of cognitive skills," and some new methods may be found. At the moment, all I can think of is this: present the students with a problem they have not encountered before (in the sense that it is conceptually different); observe (infer) how they conceptualize it, and judge what they do to solve it. It is the student's approach that is far more important than the particular solution. By observing the conceptual tools the student is using, one can usually get some inkling as to how far she or he is on the way to a viable conceptual network for the particular area. Best of all, give students a problem to which there is not yet a standard solution. I realize full well that this is not a satisfactory answer. But then, according to the statistics we read these days, the tests used at present are not satisfactory either.

Q41. ***Would a constructivist approach to science cause more students to be "better" scientists?***

A41. To answer this question, we should first have to agree what it means to be a "better" scientist. From my point of view, I would say yes but to me a scientist is better if he never forgets that science cannot reveal "objective truth," is forever fallible, and is not the most important thing in the field of human experience. A constructivist attitude is, I believe, a safeguard against the pseudo-religious faith that the search for "truth" justifies all means and can wash its hands of the possible consequences. Constructivists cannot but remain aware that they themselves are responsible for the way in which they see the world and thus for everything that may follow from their particular way of seeing. Their way of seeing is, in the last analysis, never more than one viable way of seeing.

Q42. ***In a largely positivist world, how do you suggest that constructivism can make a bridgehead in teacher education?***

A42. By having actual and prospective teachers experience that the constructivist approach works. This is not an easy task. But where it has been put into practice (at the University of Georgia, at Purdue University, at Florida State in Tallahassee, and also at my own institute at the University of Massachusetts), our initial fears and doubts have been allayed. At the beginning I was inclined to think that teachers who have taught whatever they teach for 15 or 20 years would be extremely reluctant to change their approach to teaching. But practice has proved otherwise. If one succeeds in getting teachers to make a serious effort to apply some of the constructivist methodology, even if they don't believe in it, they become enthralled after five or six weeks, not because the results in the children are so wonderful, but because the atmosphere in the classroom has changed. The relation among the children and between the children and the teacher becomes radically different—it turns into a productive relation. It is an overwhelming experience to come into a classroom of six or seven-year olds who are all intensely occupied, all doing things that are relevant to what they are supposed to be doing, and it quickly becomes evident that they are enjoying it. Most elementary school teachers find this such a relief that they are very ready to continue.

In other words, it is the success that convinces them. But there are difficulties nevertheless. It takes an enormous amount of stamina to bring about a change of mind in school boards, in principals, and in others in authority. It does not happen in a few weeks. Yet after a year or so, the radical change in the attitude of the children and, therefore in the mood of the school room, has its effect. If then it can be shown, as is the case now, that after two years the understanding generated in the constructivist classrooms is significantly greater than in traditional schools, more than half the battle has been won.

CONSTRUCTIVISM AND THE EMPIRICIST LEGACY

Michael R. Matthews

Introduction

Constructivism is one of the major influences in present day science and mathematics education. Perhaps with only slight exaggeration one participant at the 1991 US National Association for Research in Science Teaching annual conference announced 'now we are all constructivists.'

Constructivism inspires science education reform programs, is the subject of major international conferences, is the topic of hundreds of journal articles, and is the foundation of many science-teacher training programs where constructivist teaching methods are widely advocated. The theory underpins major research programs in science education, particularly those dealing with children's learning in science, misconceptions and alternative frameworks, and with curriculum development. Many have recommended the creation of science curricula on constructivist lines. And, in New Zealand for instance, one such curriculum has been developed. All of this is so well known that citation is superfluous.

Constructivism has been no less prominent in mathematics education. One well-placed writer has said: 'Anyone who observes mathematics education has to be impressed by the quite sudden eruption of "constructivism" as a central concern of so many researchers' (Davis 1990, p. 104). A glance at the Annual Conference Proceedings of the large 'International Group for the Psychology of Mathematics Education' (PME) will confirm the observation, as will a perusal of the pages of most mathematics education journals. Those not fully convinced can look at the recent anthology published by the US National Council of Teachers of Mathematics on constructivism in mathematics education (Davis, Maher & Noddings 1990). As in science education, constructivism has offered accounts of the epistemology of mathematics, the learning of mathematics, teaching strategies and curriculum development. Some of this literature is reviewed in Kilpatrick (1987), and Lerman (1989).

There is much that is laudable, insightful, and progressive about constructivist theory and practice. It is far superior to the mindless behaviourist theory of mind and learning against which Piaget, and early cognitive psychologists such as Bruner, struggled. Constructivism's stress on pupil engagement in learning, and the importance of understanding student's current conceptual schemes in order to teach fruitfully, are progressive; as is its stress on dialogue, conversation, argument, and the justification of student and teacher opinions in a social setting. Importantly constructivism's stress on understanding as the goal of science instruction is a major advance over the rote learning and mantra-like repetition of formulae that characterizes so many science classrooms—classrooms like the graduate class of which Richard Feynman said 'the students had memorized everything, but they didn't know what anything meant' (Feynman 1985, p. 212).

Given the enormous influence of constructivism, it is timely that its philosophical underpinnings be examined. This paper will contribute to such examination by trying to identify the degree to which much constructivist epistemology remains firmly, and I believe fatally, in the grip of empiricist views of knowledge development and validation. If this empiricist shadow were to be removed from constructivism then a good deal of its curriculum theorizing would be different, and a good deal of its pedagogy would also be different.

Constructivism sets out to separate itself from positivism, and empiricism, yet the separation is far from complete; the divorce ultimately does not proceed. Or to change metaphors, the situation is something like the person who renounces Christian belief, but has so absorbed Christian doctrine that, atheistic though they may be, they still interpret and act in the world in religious ways. Or again it is perhaps like the anti-Hegelian movements in German philosophy that Marx so trenchantly criticized: they decried Hegelianism, but did so using all the Hegelian categories and presuppositions.

Philosophical Background and Commitments

Constructivism, at least as it appears in science and mathematics education, has been contributed to by a number of significant intellectual movements. It had its origins in the Piagetian-led cognitive revolt against behaviourist theories of learning. Although a renowned psychologist, Piaget nevertheless understood his research program—the development of Genetic Epistemology—as primarily making a contribution to the Western philosophical tradition. Piaget's Kantian theory of mind, and his separation of the unknowable deep noumena from the knowable phenomena, are important elements for some constructivists, particularly the personal constructivists of whom von Glasersfeld is a well-known figure. (Kitchener, 1985, has provided an extensive bibliography of philosophical responses to Piaget's work.)

Constructivism has also been shaped by the widespread revolt against positivist theories of science. This revolt began in Europe in the 1930s, and spread to the Anglo-American world in the 1960s with the publication of Kuhn's The Structure of Scientific Revolutions. Subsequently most of the cherished tenets of positivism—the possibility of impartial knowledge, the possibility of sensory knowledge freed from theoretical assumptions, the uniqueness of science as a way of knowing, the inductivist view of theory development, the goal of having mind be a mirror of nature—have been severely attacked and mostly overrun. Some of the major contributors to this revolt have been Popper, Bachelard, Kuhn, Feyerabend, Toulmin, and Rorty. Despite their obvious differences on many matters, these philosophers, and numerous others, have contributed to post-positivist philosophy of science. (The first wide-ranging review of this new orthodoxy was the anthology of Suppe, 1977, another popular exposition of 'The New Philosophy of Science' was Brown, 1977. Garrison, 1986, has recently reviewed some of its main features.)

Constructivists have been influenced by the new sociology of science, in particular the social constructivism of the Edinburgh School and their so-called 'Hard Programme' in the sociology of knowledge. (Key figures in this school include Barnes, Collins, Latour, and Edge. Some philosophical responses to their work can be found in Brown 1984).

In some instances, not just post-positivist, but post-modernist views about science have been influential in constructivism. After all many, if not most, of the post-positivists held to modernist views of science: they wanted knowledge of the world, they thought that science could deliver such knowledge, they just denied that positivism gave an adequate account of the processes and outcomes of the scientific endeavor. More recently in the work of Foucault, Derrida, and other post-modernists, such assumptions have been challenged.

Other intellectual traditions have contributed to the heterogeneous doctrine we call constructivism: the personal-construct psychology of Kelly has been utilized; and of special significance for social constructivists has been the theory of language acquisition originally developed by Vygotsky in Russia in the 1930s. It is not my intention here to chart all these influences, I only wish to establish that constructivism is more than a psychological theory, that it has a significant philosophical component.

Constructivism in science and mathematics education is not just a theory of human concept formation and learning, nor just a theory of the way science develops and is validated, it is perhaps preeminently a theory of education (to use the title of one prominent constructivist's book—Novak, 1977). It is a view about how teaching should proceed, how children should be treated, how classrooms should be organized, how curricula should be developed and implemented, and sometimes even a view about the purposes of schooling.

Given these contributors to its lineage, it is not surprising that constructivism fairly bristles with philosophical commitments. It explicitly assumes and champions positions in the philosophy of science, the philosophy of mind, and the philosophy of education. It is a comprehensive doctrine; at once a theory of science, of human learning, and of teaching.

The range of constructivist concerns can be seen in the following sub-headings contained in a recent science education article: 'A Constructivist View of Learning,' 'A Constructivist View of Teaching,' 'A View of Science,' 'Aims of Science Education,' 'A Constructivist View of Curriculum,' 'A Constructivist View of Curriculum Development' (Bell 1991).

In this paper I wish to concentrate upon the epistemological commitments of constructivism. Some extracts from constructivist writings can give a sense of the pervasiveness of these commitments:

> such [constructivist] approaches view knowledge as personally and socially constructed, rather than 'objective' and revealed; theories as provisional, not absolute (Millar & Driver 1987, p. 57).
>
> Science as knowledge is an intellectual construct, and what are referred to as the laws of nature are merely the result of this human activity. Nature as such does not have laws (Nadeau & Desautels 1984, p. 19).
>
> The fact that scientific knowledge enables us to cope does not justify the belief

> that scientific knowledge provides a picture of the world that corresponds to an absolute reality (von Glasersfeld 1989, p. 135).
>
> Although we may assume the existence of an external world we do not have direct access to it; science as public knowledge is not so much a discovery as a carefully checked construction (Driver & Oldham 1986, p. 109).
>
> Scientific ideas and theories not only result from the interaction of individuals with phenomena but also pass through a complex process involving communication and checking through major social institutions of science before being validated by the scientific community (Driver 1989, p. 85).
>
> Put into simple terms, constructivism can be described as essentially a theory about the limits of human knowledge, a belief that all knowledge is necessarily a product of our own cognitive acts. We can have no direct or unmediated knowledge of any external or objective reality. We construct our understanding through our experiences, and the character of our experience is influenced profoundly by our cognitive lens (Confrey 1990, p. 108).
>
> Scientific knowledge is invented in order to make sense of observations which are themselves theory-laden. There is no great book of nature that can be consulted in order to check if the models or theories correspond to an ontological reality (Desautels & Larochelle 1990, p. 236).
>
> Constructivism is, logically, a post-epistemological position. The standard questions of epistemology cannot be answered—or even reasonably asked—from this perspective. Its premises suggest, rather, abandonment of traditional epistemological language (Noddings 1990, p. 18).

However epistemology, even when abandoned, is important to constructivism, indeed I believe it drives the rest of constructivist educational theory and practice.

As previously indicated, constructivism is a heterogeneous movement. One recent review has identified at least the following varieties: contextual, dialectical, empirical, information-processing, methodological, moderate, Piagetian, post-epistemological, pragmatic, radical, realist, social and socio-historical (Good, Wandersee & St Julien 1991). To this list could be added humanistic (Cheung & Taylor 1991). One widely acknowledged division is between personal and social constructivists. I believe that empiricism casts a shadow over all of these sub-schools, but admittedly some may be more in shadow than others.

Constructivist ontology is largely realist—the world although unknowable does actually exist apart from our thinking about it. But some constructivists are idealist—the world is created by human thought and dependent upon such thought. Ernst von Glasersfeld's radical constructivism is the best known idealist variant. He says:

> the realist believes his constructs to be a replica or reflection of independently existing structures, while the constructivist remains aware of the experiencer's role as originator of all structures...for the constructivist there are no structures other than those which the knower constitutes by his very own activity of coordination of experiential particles (von Glasersfeld 1987, p. 104).

His idealism is admittedly difficult to pin down. At one point he does say that 'Superficial or emotionally distracted readers of the constructivist literature have frequently interpreted this stance as a denial of "reality"' (von Glasersfeld 1991, p. xv) but immediately after this disclaimer he goes on to say:

> the fact that we *do* agree on certain things and that we *can* communicate does not prove that what we experience has *objective reality* in itself. If two people or even a whole society of people look through distorting lenses and agree on what they see, this does not make what they see any more real—it merely means that on the basis of such agreements they can build up a consensus in certain areas of their subjective experiential worlds.

Grayson Wheatley is a still more ambiguous idealist, but at one point idealism shines through clearly when he says:

> From a constructivist perspective, knowledge originates in the learner's activity performed on *objects*. But objects do not lie around ready made in the world but are mental constructs (Wheatley 1991, p. 10).

Although there are variations in constructivist epistemology, the common thread is that it is subject-centered, experience-based, and relativist. But its relativism needs to be distinguished from other relativisms in which the goal of science as a search for truth about the world is accepted, and it is then asserted that we cannot know of different accounts which one is actually true or better. In contrast for most constructivists our knowledge does not tell us about the world at all, it tells us about our experiences, and how they are best organized.

Lerman (1989), following Kilpatrick (1987), suggests that the core epistemological theses of constructivism are:

(1) Knowledge is actively constructed by the cognizing subject, not passively received from the environment.

(2) Coming to know is an adaptive process that organizes one's experiential world; it does not discover an independent, pre-existing world outside the mind of the knower.

The first thesis (1) is a psychological claim, the second (2) is an epistemological claim. Lerman believes (1) implies (2).

Wheatley offers a nearly identical summation of the epistemological core of constructivism. He says:

> The theory of constructivism rests on two main principles... Principle one states that knowledge is not passively received, but is actively built up by the cognizing subject... Principle two states that the function of cognition is adaptive and serves the organization of the experiential world, not the discovery of ontological reality... Thus we do not find truth but construct viable explanations of our experiences (Wheatley 1991, p. 10).

Constructivism as a form of Empiricism

Given the influence of constructivism, surprisingly few critiques of its epistemology, or more broadly its philosophy, have appeared in educational literature. Strike (1987) provides one noteworthy criticism. There is some skirmishing between mainstream Piagetian personal constructivists and Vygotskian social constructivists; and there have been some refinements and adjustments around the epistemological edges, for example by Kilpatrick (1987), Lerman (1989) and Goldin (1990)—but all of these can be regarded as domestic disputes.

It is as if a period of Kuhnian normal science has descended upon the science and mathematics education communities. The President of the US National Association for

Research in Science Teaching suggests as much: 'A unification of thinking, research, curriculum development, and teacher education appears to now be occurring under the theme of constructivism [there is] a lack of polarized debate' (Yeany 1991, p. 1). A recent long and detailed criticism of von Glasersfeld's philosophy by Suchting (1992) might mark a change in this situation.

My criticism of constructivist epistemology is, in brief, that constructivism maintains the Aristotelian-empiricist epistemological paradigm, and by correctly pointing to a major error in empiricist assumptions, it then swings to a relativist epistemology without abandoning the paradigm itself. The relativist conclusion only follows within the empiricist paradigm. If this paradigm is rejected, no such relativist epistemological conclusions follow, and certainly no idealist ontological conclusions follow.

To understand my argument it is necessary to allude to the Aristotelian conception of knowledge which in one guise or another has so dominated the history of epistemological reflection.

Aristotle assumed a world independent of people populated by individual substances that formed classes or species in virtue of their essences or forms. To know something was to know its essence, or nature. Knowers were in two important senses passive in the Aristotelian theory of knowledge: first, the world imprinted empirical knowledge on the subject through ordinary unmediated sense-perception, this had to be unmediated in order for the object of knowledge to be known as it actually was; second, the knower could not interfere with what he or she was hoping to know. Aristotle held a non-interference epistemological policy. Once this stand-off between the knower and the known was achieved, this ideal epistemological situation—neither the knower nor the known being interfered with—then the rational faculty, nous, of the knower by a process of induction or epagoge intuited the universal forms inherent in the known.

As Dewey and Popper have observed, Aristotle basically held an empiricist, individualistic, reflective or correspondence—'Spectator'—theory of knowledge: knowledge was something generated by, and residing inside, an observer. The essence of objects was displayed in their natural motions; violent motion was the result of interference with this natural motion and it was thus of no interest to science, indeed it was an obstacle to science. To the extent that the world or object was interfered with then knowledge was thwarted. To the extent that the subject was active in knowledge acquisition, then knowledge was compromised. Thus there was no experimental tradition in Aristotelian science, it was a tradition of observing nature in its natural state; a tradition that placed great emphasis on the eye, not the hand.

Observation, vision, and looking, were the key elements in the subject side of Aristotle's epistemological equation, on the other object side the key element was undisturbed nature. Knowledge was obtained when the subject side accurately mirrored the object side. Aristotle's epistemological dictum was: 'to say of what is that it is, and of what is not that it is not, is true.' With the Scholastics this becomes the doctrine that truth is the adequation (or correspondence) of things and the intellect.

Despite their anti-Aristotelian rhetoric, the British empiricists of the seventeenth century—Bacon, Locke, Berkeley, and Hume—maintained in all essentials this Aristotelian paradigm of knowledge. Knowledge was an individual matter, it was grounded in experience or sense impressions, its criterion was adequacy between thought and reality or the correspondence of ideas with objects.

Any epistemology which formulates the problem of knowledge in terms of a subject looking at an object and asking how well what they see reflects the nature or essence of the object is quintessentially Aristotelian, or more generally empiricist. Aristotelians were direct realists about perception—that is, the objects of perception were material bodies, what we looked at were bodies in the world. Later empiricists, Locke, Hume and Berkeley, were largely indirect realists—that is, the objects of perception were sense impressions generated, it was supposed, by material objects.

Locke, an avowed opponent of Aristotle, in his 1690 *Essay* puts the matter this way: 'The mind, in all its thoughts and reasonings, hath no other immediate object but its own ideas, which it alone does or can contemplate.' Hume in the opening pages of his 1739 *Treatise* takes the empiricist discussion further by dividing the contents of the mind into *impressions* and *ideas*. The former are, roughly, sensations, and are thought to give rise to the latter which are 'faint images of these in thinking and reasoning.' At least this is the case for *simple ideas* where 'every simple impression has a correspondent idea.' Thus to 'give a child an idea of orange...I present the objects, or in other words, convey to him these impressions.'

As is well known Locke's formulation of the problem of knowledge was used by Berkeley to support idealism and relativism. Berkeley's argument in his *Treatise* was simple but devastating: 'As for our senses, by them we have the knowledge *only of our sensations*, ideas, or those things that are immediately perceived by sense, call them what you will: but they do not inform us that things exist without the mind, or unperceived.'

Variations of this classical empiricist account of the origin of ideas recur in contemporary constructivist writings. Consider Osborne and Wittrock's influential 1985 generative learning model. There they say that the key postulates of the model include:

> •The learner's existing ideas influence what use is made of the senses and in this way the brain can be said to actively select sensory input.
>
> •Learner's existing ideas will influence what sensory input is attended to and what is ignored.
>
> •The input selected or attended to by the learner, or itself, has no inherent meaning.

And they proceed to discuss constructivist teaching in terms of 'sensory input' provided by the teacher. They acknowledge that the construction or generation of complex meanings is not straightforward. This has to be done by the students 'producing their own sensory inputs'—tables, diagrams, flow charts, etc. It is claimed that the science classroom is a particularly appropriate site for testing the generative learning model because:

> The sensory input can be clearly specified (e.g., reading a scientific text), or consist of a full range of sensory input (e.g., the science laboratory experiment), depending on the nature of the particular research study.

In passing one might note here that the supposed clarity with which the sensory inputs associated with reading a scientific text can be specified is somewhat problematic. Here the assumed separation of meaning from stimulation is most tendentious, as was shown by Gibson in his classic of the 1950s, 'What is a Stimulus?'

It is not coincidental that modern constructivists once having formulated the epistemological problem in Aristotelian-Lockean terms then endorse versions of Berkeley's savage critique of it and end up with relativism, and for the more consistent, idealism. It is not for nothing that von Glasersfeld says that 1710 was a very good year: being the year of publication of both Vico's major work, On the Most Ancient Wisdom of the Italians, and Berkeley's Treatise.

That a leading constructivist identifies Vico and Berkeley as the founders of constructivism and lauds their philosophy is indicative of the ambiguous relationship between constructivism and modern science, including science education: Vico and Berkeley were the avowed opponents of the Scientific Revolution and of the work of Galileo, Newton, and others. Both of them were defenders of orthodox Christianity against the perceived irreligious philosophy of the new science. Berkeley's 1710 Treatise was subtitled 'Wherein the Chief Causes of Error and Difficulty in the Sciences, With the Grounds of Scepticism, Atheism, and Irreligion are Inquired Into.' That so many today are defending science education with the doctrines of two deeply anti-scientific thinkers is cause for some reflection if not misgivings.

Within this paradigm, framework, or problematic, the possibility of knowledge weakened once it was pointed out that the mind is active in cognition, not just in intuiting the forms, but in structuring the experiences that the senses provide; and the possibility of knowledge of an external world evaporated once it was claimed that the immediate objects of the intellectual faculty were sense impressions rather than nature itself. Nature, or in Kant's terms, the thing-in-itself, became unknowable, because we only ever see it through distorting lens, and there is no privileged position from which to check the adequacy, or correspondence, of thought to reality.

Constructivism's acceptance of the fundamentals of the Aristotelian-empiricist epistemological problematic is indicated when von Glasersfeld speaks of 'looking through distorting lens and [agreeing] on what they see,' and Confrey speaks of 'cognitive lens,' and Desautels and Larochelle write of 'making sense of observations which are themselves theory-laden'; and when numerous others have recourse to this looking/seeing/observing vocabulary for stating the problem of knowledge, or when use is made of the Kuhnian/Hansonian ambiguous or hidden figures to establish facts about the theory dependence of observation, or when gestalt-switch terminology is used to describe scientific revolutions. Whether subjects are seeing through lens clearly or darkly, it is the metaphor of seeing through lens which signals commitment to an empiricist theory of knowledge.

The one-step argument from the psychological premise 'the mind is active in knowledge acquisition' to the epistemological conclusion 'we cannot know reality' is endemic in constructivist writing. Lerman speaks for many when he says of the two theses—(1) and (2) above—that 'the connections between hypothesis (1) and (2) seem to be quite strong' (1989, p. 212).

But this conclusion only follows on the assumption that Aristotle had in fact correctly delineated the problem of knowledge; the conclusion only follows if the terms of the problem as stated by Aristotle are accepted. If this assumption is rejected then none of the skeptical conclusions of constructivism follow.

The Scientific-Epistemological Revolution

Although presented here in a condensed and largely suggestive manner, I will use the example of the scientific and epistemological revolution of the seventeenth century to illustrate the inadequacy of empiricism generally, and constructivism specifically, to grasp, let alone solve, the epistemological problem presented by modern science. The example also allows for some of the powerful, but partial, insights of constructivism to be demonstrated.

The scientific revolution associated with Galileo effectively destroyed the Aristotelian epistemological paradigm. Unfortunately the seventeenth-century empiricist champions of the new science saw only Aristotelian science as being refuted in the work of Galileo and Newton. They interpreted this scientific revolution in terms of the old epistemology—saying that the new science gave a better fit between thought and reality, that it was more adequate to experience, that the scientist-philosophers were better thinkers than their predecessors, etc. This mismatch between the methods of the New Science, and what the philosophical defenders of the New Science thought its methods were, has been much commented upon; perhaps no where better than by Mittelstrass (1972).

Galileo's accomplishment is to make the important distinction between the *theoretical* object of science, which is a system of mathematically expressed definitions, principles, concepts, and relations; and the *real* objects of science, which are the materials, events, objects in the world that are grasped, described, and, by suitable instrumentation and experimentation, manipulated by scientists, and others, in accordance with the former. These theoretical objects are not Lockean ideas in contrast to material bodies; neither are they Humean ideas arising automatically from sense impressions.

Galileo's account of pendulum motion well illustrates the difference between the theoretical objects of Galileo's science and the real objects of the world. (On this matter see Matthews 1987, 1992).

Mythology not withstanding Galileo did not look carefully at pendulums swinging and then derive his law of isochronic motion. In fact his Aristotelian patron, and on this occasion opponent, Guidibaldo del Monte did just this and not surprisingly thought that Galileo had taken leave of his senses when he formulated his law. Real pendulums in no way behave isochronically—if they did, an immediate consequence of Galileo's law would make pendulums perpetual mobiles, which clearly they are not. Rather Galileo began with a collection of *concepts* in dynamics, with a *mathematical* means of depiction and analysis that dealt with *abstracted* circumstances where many features of real pendulums were ignored—their color, massiveness, etc., and with *idealized* situations where the bob was regarded as a point mass, where there was no mechanical friction at the fulcrum nor air resistance in the swing, and where the string had no weight and hence did not damp the motion, etc. When the real object—the briefly swinging, non-isochronic, material pendulum—was described in terms provided by the theoretical object, and depicted geometrically, it was then in a suitable form for analysis by Galilean science. The result was the law of isochronic motion. Del Monte said that this was a world on paper, and that science was to be about the real world. Crucially Galileo's science was about the real world, he was not a Platonist or idealist, but it was not immediately about this world as naturally disposed: it was mediated by the conceptual-theoretical object which clearly did not correspond to the real world.

Galilean science applied to a highly constrained experimental situation formed only on the design provided by the theoretical object. The real was altered, that is refined and perfected, to correspond with the theoretical, not the other way around as del Monte the Aristotelian, and more generally empiricists would insist in their claim that theory should be brought into line with reality. That the theoretical object did not correspond with the real was beside the point. Subsequently an interplay of theory—abandoning for instance Galileo's view that the circle was the brachistochrone, or curve of quickest descent—and experimental refinement lead to Huyghen's improvement of the law of pendulum motion.

Seeing or observing things differently was just not important in the debate between Galileo and del Monte. Galileo did not see point masses at the end of weightless strings, nor did he see continuous motion at the zenith of the swing. He saw exactly what del Monte saw—chandeliers at the end of chains, lead balls on the end of wire, pendulums which momentarily stopped at their zenith, and pendulums which came to a halt within about two-dozen oscillations, etc. Observation—whether theory dependent or not—was barely relevant to the dispute. Galileo described things differently, he did not see things differently. The new descriptions did not come from looking—after all one of the greatest observers of all time, Leonardo da Vinci, had long looked at pendulums without describing them in the way Galileo did—the descriptions came from an intellectually constructed theoretical object. Once described mathematically these statements were able to be worked upon by his theory of mechanics.

As one commentator has said: 'observation and experience, in the sense of brute, commonsense experience, did not play a major role—or if it did, it was a negative one, the role of obstacle, in the foundation of modern science' (Koyre 1968, p. 18).

Constructive Criticism

Constructivism is correct in stressing the inventive, humane, culturally and temporally dependent aspects of creating the theoretical objects of science—but none of this bears, of itself, upon truth. Further, although constructivists stress this creative aspect of knowledge production, they seldom extend the analysis. In particular for personal constructivists, their model of creation is a sort of individual, cottage-industry, model. It is the personal, Robinson Crusoe, model of knowledge construction that leaves aside the necessary social and communitarian dimensions of cognition—the dimensions that inclined Freire, following Hegel, to say that 'it is the "we think" that determines the "I think" and not the other way around.'

Wheatley, quoted above, is partially correct when he says that the objects of science do not just lie around. Where he and most constructivists, in company with Aristotle and the empiricists, go wrong is in failing to distinguish the theoretical objects of science, which do not lie around, from the real objects of science, which do lie around and fall on people's heads.

Importantly even the theoretical objects, once produced, have a reality even if they are not exactly lying around. Newton's mechanics, Darwin's evolutionary theory, and Mendel's genetics all exist and can be apprehended by, and taught to, subsequent thinkers—it is just that they are not to be confused with falling apples, grazing animals, or fields of peas.

Karl Popper well expressed the objectivist theory of knowledge that can be counterposed to personal constructivism:

> My...thesis involves the existence of two different senses of knowledge or of thought: (1) knowledge or thought in the subjective sense, consisting of a state of mind or consciousness or a disposition to behave or act, and (2) knowledge or thought in an objective sense, consisting of problems, theories, and arguments as such. Knowledge in this objective sense is totally independent of anybody's claim to know; it is also independent of anybody's belief, or disposition to assent; or to assert or to act (Popper 1979, p. 108).

The idea of theoretical production is much more verdant than is usually realized. Marx said that people create history, but they do not create it as they choose; they create it in definite circumstances over which they have little control. As with material production, so also theoretical production depends upon raw materials (words, concepts, language, formulae), means of production (mathematical systems, logic, computational devices, instruments), productive relations (ways in which science is organized and funded, ownership of intellectual property, educational arrangements), and finally the social and legal context of the productive activity (religious constraints, economic pressures, cultural values). Knowledge is a form of theoretical production and can usefully be analyzed as such.

Empiricism, and constructivism, conceives the enterprise of science in terms of individuals looking at the world and trying to ascertain whether their ideas, concepts, and conceptualizations make sense. In the words of the New Zealand Draft 1-5 Syllabus: 'Science is about people exploring and investigating their biological, physical, and technological worlds, and making sense of them in logical and creative ways' (p.5).

Objectivist epistemology distinguishes between the raw material and events of the world (the real objects of science), the theoretical structures and concepts of science, the material and events as described by the theory (the theoretical objects of science), and the experimental and technical procedures of science. Individual understanding and conceptualization is parasitic upon this extra-individual scientific domain. Science education can then be conceived in terms of the appropriate introduction of individuals into this world of concepts, understandings, techniques, and community standards.

Once the distinction is made, and knowledge is recognized as a process of intellectual production working with real objects that have been described, apprehended, or incorporated by a theoretical object, then the interesting epistemological tasks of evaluating different modes of knowledge production—in terms of fecundity, simplicity, utility, truthfulness—can be commenced. Relativism short-circuits all of this.

Constructivists, as with turn-of-the-century positivists such as Mach, Poincare, and Duhem, once they roll back our knowledge to knowledge of sensory input, then are inexorably led to speak of the aims of science as being the 'making sense of our world,' or for the more consistent, 'making sense of our sense impressions.' Alternatively this view leads to a purely instrumentalist interpretation of scientific theory. Both alternatives are in contrast to realism where the aim of science is the understanding or explanation of our world. Science may not achieve the aim, but there is agreement that science is the attempt to comprehend something outside ourselves.

Once this positivist limitation on scientific knowledge is accepted, then constructivist recommendations for the science curriculum inexorably move to formulations such as students should 'make sense of the world around them,' or 'make sense of the living world.' This talk of 'making sense' is quintessentially empiricist. It is also fraught with grave educational and cultural implications. It leads immediately, as it has historically done, to relativisms of all kind.

It is notorious that people have for centuries thought that the grossest injustices, and the greatest evils, have all made sense. Likewise most scientific advances have entailed commitment to propositions that did not only defy sense—Galileo's point masses and colorless bodies, Newton's inertial systems, Darwin's gradualist evolutionary assumptions, Einstein's mass-energy equivalence, etc.—but also defied accepted scientific, social, cultural, or religious norms. 'Making sense' of our sensory inputs hardly seems sufficient warrant to maintain the scientific enterprise; and in a science classroom it hardly seems sufficient warrant for the teacher to disturb deeply ingrained and important beliefs of children. Finding out the truth might provide such warrant.

And if, as some constructivists say, we merely start with what makes sense to students, and then proceed to encourage them along to the current scientific view, despite intermediate false views seeming to make sense to them, then the question arises as to why do we do this? What is it that we think is important about science that warrants this interference with ingrained and comfortable belief? It can only be that science makes more sense. But the reason it makes more sense cannot be spelled out in terms of sensory inputs. This whole formulation of the issue is simply mistaken.

Von Glasersfeld says often that the orthodox epistemological problem has to be abandoned, and that correspondence as a mark of truth needs to be rejected. This is because on his terms we cannot see reality, we only have our sensations to reflect upon, and so we are never able to judge correspondence between our ideas and the world (a restatement of Berkeley's argument). He sometimes replaces correspondence with pragmatism, and in other places with coherence among experience. My suggestion is that this is a case of new wine in the old Aristotelian bottle. This leads him, and Noddings for instance, to speak of constructivism as being post-epistemological. His original subject-observing-an-object formulation does lead to an impasse, and one appreciates why he wants to abandon epistemology. But the mistake is to give up wine drinking because all the contents of one vintage are undrinkable; there are other vintages which allow epistemological embibing to proceed with interest, enjoyment, and profit.

References

Bell, B. (1990). Draft forms 1-5 syllabus. Wellington: Ministry of Education.

Bell, B. (1991). A constructivist view of learning and the draft forms 1-5 science syllabus. SAME Papers, Hamiliton New Zealand: University of Waikato.

Brown, H.I. (1977). Perception, theory and commitment: the new philosophy of science. Chicago: University of Chicago Press.

Brown, J.R. (ed.). (1984). Scientific rationality: the sociological turn. Dordrecht: Reidel.

Cheung, K.C. & R. Taylor. (1991). Towards a humanistic constructivist model of science learning: changing perspectives and research implications. Journal of Curriculum Studies, 23 (1).

Confrey, J. (1990). What constructivism implies for teaching. In R.B. Davis, C.A. Maher, & N. Noddings (eds.), Constructivist views on the teaching and learning of mathematics. Reston, VA: National Council of Teachers of Mathematics, 107-122.

Davis, R.B. (1990). Discovery learning and constructivism. In R.B. Davis, C.A. Maher, & N. Noddings (eds.), Constructivist views on the teaching and learning of mathematics. Reston, VA: NCTM, 93-106.

Davis, R.B., Maher, C.A., & N. Noddings (eds.). (1990), Constructivist views on the teaching and learning of mathematics. Reston, VA: National Council of Teachers of Mathematics.

Desautels, J. & M. Larochelle. (1990). A constructivist pedagogical strategy: the epistemological disturbance (experiment and preliminary results). In D.E. Herget (ed.), More history and philosophy of science in science teaching. Tallahassee, FL: Florida State University.

Driver, R. (1989). The construction of scientific knowledge in school classrooms. In R. Millar (ed.), Doing science: images of science in science education. Lewes, East Sussex: Falmar Press, 83-106.

Driver, R., & V. Oldham. (1986). A constructivist approach to curriculum development in science. Studies in Science Education, 13, 105-122.

Feynman, R. (1986). Surely you're joking Mr. Feynman. New York: Bantam Books.

Garrison, J.W. (1986). Some principles of postpositivist philosophy of science. Educational Researcher, 15 (9), 12-18.

Goldin, G. (1990). Epistemology, constructivism, and discovery learning in mathematics. In R.B. Davis, C. Maher, & N. Noddings (eds.), Constructivist views on the teaching and learning of mathematics. Reston, VA: NCTM.

Good, R., Wandersee, and St Julien. (1991). Cautionary notes on constructivism. Unpublished paper.

Kilpatrick, J. (1987). What constructivism might be in mathematics education. In J.C. Bergeron, N. Herscovics, & C. Keiran (eds.), Psychology of mathematics education. Montreal: Proceedings of the Eleventh International Conference, 3-27.

Kitchener, R.F. (1985). A bibliography of philosophical work on Piaget. Synthese, 65 (1), 139-152.

Koyre, A. (1968). Metaphysics and measurement. Cambridge, MA: Harvard University Press.

Lerman, S. (1989). Constructivism, mathematics, and mathematics education. Educational Studies in Mathematics, 20, 211-223.

Matthews, M.R. (1987). Galileo's pendulum & the objects of science. Philosophy of Education, 309-319.

Matthews, M.R. (1990). Galileo and pendulum motion: a case for history and philosophy in the science classroom. The Australian Science Teachers Journal, 36 (1), 7-13.

Millar, R. & R. Driver. (1987). Beyond processes. Studies in Science Education, 14, 33-62.

Nadeau, R. & J. Desautels. (1984). Epistemology and the teaching of science. Ottawa: Science Council of Canada.

Noddings, N. (1990). Constructivism in mathematics education. In R.B. Davis, C. Maher, & N. Noddings (eds.), Constructivist views on the teaching and learning of mathematics. Reston, VA: NCTM, 7-18.

Novak, J. (1977). A theory of education. Cornell: Cornell University Press.

Popper, K.R. (1972). Objective knowledge: an evolutionary approach. Oxford: Oxford University Press.

Strike, K.A. (1987). Towards a coherent constructivism. In J.D. Novak (ed.), Misconceptions and educational strategies. Cornell University: Education Department Vol. 1, 481-489.

Suchting, W.A. (1992). Constructivism deconstructed. Science and Education, 1 (3), 223-254.

Suppe, F. (ed.). (1977). The structure of scientific theories. Urbana, IL: University of Illinois Press.

von Glasersfeld, E. (1987). The construction of knowledge. Seaside, CA: Intersystems Publications.

von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. Synthese, 80 (1), 121-140.

von Glasersfeld, E. (1991). Introduction. In E. von Glasersfeld (ed.), Radical constructivism in mathematics education. Dordrecht: Kluwer, xiii-xx.

Wheatley, G.H. (1991). Constructivist perspectives on science and mathematics learning. Science Education, 75 (1), 9-22.

Yeany, R.H. (1991). A unifying theme in science education? NARST News, 33 (2), 1-3.

HOW DO STUDENTS' VIEWS OF SCIENCE INFLUENCE KNOWLEDGE INTEGRATION?

Nancy Butler Songer and Marcia C. Linn

Do scientists ever disagree with each other? Will scientific principles found in classroom textbooks always be true? How do scientists verify new ideas? Which is better: to memorize facts or try to understand complicated material when learning science? Student answers to these questions reveal a great deal about their views of science, their strategies for learning science, and their accomplishments in science courses.

In this article we characterize the beliefs that students hold about science and about the learning of science material. We analyze the relationship between these beliefs and the integration of scientific knowledge in the domain of thermodynamics. Finally, we determine how students respond to concrete, as opposed to abstract, loci for knowledge integration and assess whether these preferences are related to beliefs about science.

Beliefs about Science

What do we mean by "student beliefs about science?" We sought first to characterize students' views of scientific knowledge and of scientists, and second to characterize their views about how science knowledge should be acquired.

What is a productive view of science knowledge? Historians of science report that science knowledge is socially constructed. The personalities of those involved have an impact on the pace of scientific advance as well as on the selection of the research agenda. This point is dramatically illustrated in the controversy between Galileo and the Catholic church (Brecht, 1980). Historians and philosophers of science also point out that science progresses by fits and starts, with either evolutionary or revolutionary changes in perspective (Kuhn, 1970). In addition, science knowledge is controversial, especially at the time of its discovery. Many current debates about treatments of AIDS, prediction of earthquakes, or extinction of dinosaurs illustrate the controversial nature of science. Finally, scientific knowledge is relevant to the lives and problems of individuals as well as societies. Fundamentally, historians, philosophers, and scientists

---

This article first appeared in the ***Journal of Research In Science Teaching*** *(1991), Volume 28, Number 9, pp. 761-784. Copyright © 1991 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

agree that science is a dynamic enterprise involving conjecture, debate, verification, reanalysis, political influence, and a certain amount of good fortune.

The view of science communicated by historians and philosophers of science, as well as by many scientists themselves, differs markedly from the perspective one might develop from reading a typical science textbook (see also Reif & Larkin, 1991). Students could develop views of science that are quite divergent from those held by historians, philosophers, and expert scientists, because textbooks often communicate the results of this process, but give little indication of the process itself.

Besides studying beliefs about scientific advance, we also examine beliefs about learning science. These two types of beliefs are, of course, closely entwined, because conjectures about what one is learning are likely to influence the methods one uses for acquiring this knowledge. Many students report that they prefer to memorize scientific information rather than trying to understand it. Often the tactic of memorizing information is preferred because it is more efficient or better attuned to the assessment techniques found in many courses. Yet, if science is a dynamic, socially constructed set of ideas, then memorization is unlikely to yield understanding of how knowledge develops or of the nature of science itself.

How might beliefs about science develop? Elsewhere we have described development of scientific conceptions as moving from action knowledge to intuitive beliefs, and ultimately to scientific principles (Linn & Songer, in press a; Linn, Songer, Lewis, & Stern, in press). This framework can also be applied to the development of beliefs about science. Furthermore, other researchers have identified similar developmental progressions to explain how students come to understand scientific models (Grosslight, Unger, & Jay, 1991) and relationships among gears (Metz, 1991).

We describe knowledge acquisition as beginning with action and observation. We refer to the initial, unreflective ideas that students develop as "action knowledge." In the case of beliefs about science, action knowledge may include such beliefs as "science doesn't make sense," or "science is what the textbook says," or "science never changes."

We argue that action knowledge is combined using the process of reflective abstraction (Piaget, 1952) into "intuitive conceptions." These conceptions are the conjectures students make to explain observations or experiences. For example, students may develop intuitions such as "scientists use words that other people can't understand," or "scientists' theories apply to laboratory experiments, not everyday life." Students may develop contradictory intuitions. Thus, students may believe that scientists agree with each other and point to textbook statements along the lines of "scientists have found that matter is made up of particles." They may also believe that scientists disagree with each other and cite newspaper or television accounts of scientists arguing about cold fusion or treatments for cancer. These intuitive ideas may seem contradictory to an outside observer, yet understandable from the standpoint of the student. For example, a student may resolve this dilemma by saying that treatments for diseases are not part of "science." Similarly, in the area of thermodynamics, we find that students often hold separate beliefs about the nature of heating and the nature of cooling, believing, for example, that wrapping a cold drink in aluminum foil will keep it cold, and that an aluminum foil jacket would not keep someone warm. The development of scientific knowledge involves integrating beliefs about individual events, and recognizing that seemingly unrelated events are in fact explained by the same set of ideas.

We assert that effective instruction can help students sift through their intuitions and combine predictive intuitions into what we call "scientific principles." Principles are abstract general rules that explain what might be viewed as unrelated events. Examples include, "energy is conserved" or "objects in motion tend to remain in motion." Principles governing views of science might include "science proceeds by fits and starts" or "science is socially constructed."

In this article we analyze students' beliefs about science in terms of this developmental progression. We identify students in terms of whether they have action knowledge or intuitive beliefs about science. In addition, we characterize their intuitions as reflecting a dynamic or static view of science.

Knowledge Integration

What do we mean by knowledge integration? We have argued that as students develop understanding of science, they organize action knowledge into intuitive beliefs and ultimately into principles. This process of organizing information into broader categories and into more widely applicable ideas results in knowledge integration. Our previous work indicates that expert teachers help students integrate such things as their understanding of heating with that of cooling (Linn and Songer, in press b). In addition, the process of knowledge integration involves distinguishing ideas that may seem similar at the onset to "make sense" of events. In thermodynamics, students often start with the notion that heat and temperature are the same phenomenon. Expert teachers can help them distinguish heat and temperature in order to develop a more predictive model of thermodynamics.

Songer (1989) reports on the effectiveness of instruction using a concrete integration aid. Her instruction helped students recognize that, for example, Styrofoam is a better insulator than metal, and that wool is better than metal as well. She then encouraged the students to organize these disparate pieces of knowledge along a continuum line with insulators at one end and conductors at the other. These comparisons were qualitative. Obviously, the thickness of the material or the conditions under which the material was applied could also influence the effectiveness of the material as an insulator or a conductor. Her training was very successful. Students who used the continuum line were better at understanding the nature of conduction and insulation than those who carried out experiments to investigate various materials and their effect on insulation but did not have the benefit of a continuum line as an integration aid (Songer, 1989).

Thus, by integration we mean the synthesizing of ideas into a cohesive and coordinated whole. Sometimes integration means recognizing that two seemingly different processes are really explained by the same underlying principles; at other times, in order to integrate a set of experiences, it may be necessary to differentiate concepts such as heat and temperature which had previously been seen as identical.

How do students develop integrated understanding? Songer's success with the continuum line suggests the importance of integration aids. In this study we explore two additional integration aids: pragmatic principles about thermodynamics, and prototypic thermodynamic events. By pragmatic principles we mean abstract principles that summarize thermodynamic experiments but are accessible to students. We distinguish pragmatic principles from expert principles and argue that pragmatic principles are more commonly qualitative than quantitative and more likely to be macroscopic than microscopic. In the case of thermodynamics, we (Linn et al., in press) identified pragmatic principles based on the heat-flow view of thermodynamics popular among scientists in the 1850s. Our version of the model includes the caveat that heat does not have mass, but in other respects is quite similar to a 1850s view of thermodynamics (see principles in Figure 1). In contrast, the view of thermodynamics based on molecular-kinetic theory is, in our estimation, abstract and microscopic, and therefore very difficult for eighth graders to integrate with their experiences and understandings. Our conjecture is that by offering students pragmatic models in eighth grade, we pave the way for their understanding of expert models when they return to this topic.

By prototypes, we mean familiar situations that illustrate a class of scientific events and for which most students have accurate predictions. For example, students can accurately predict what will happen if they stir hot liquid with a metal spoon and a wooden spoon. They point out that the metal spoon will get hot, but that a wooden spoon will remain cooler. Instruction can build on this prototypic experience to help students aggregate situations where different materials are in contact with the same heat source. The students are encouraged to recognize that the rate of heat flow depends on the material. Our use of prototypes to help students understand thermodynamics is consistent with Clement's (Clement, Brown, & Zietsman, 1989) call for bridging analogies and with Minstrell's (1982) benchmark experiments. Prototypes are more concrete than principles. Students whose understanding is quite disconnected, who have more action knowledge than intuitive beliefs, are likely to find prototypes more useful than principles. In contrast, students who already have well-developed intuitive beliefs are likely to benefit from scientific principles. Principles allow students to gain a more general and abstract view of the material they are learning.

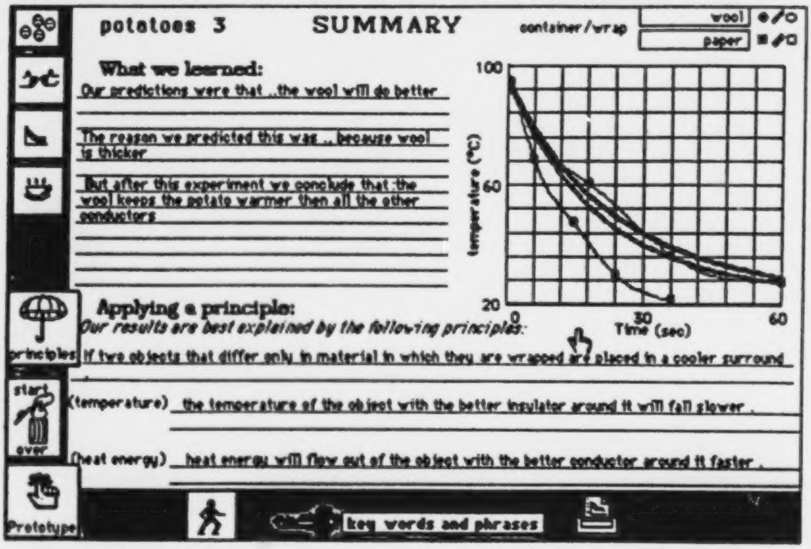
In this study, we contrast principles and prototypes as the locus for knowledge integration and report on students' reactions on these different forms of instruction. In addition, we relate students' efforts at knowledge integration to their beliefs about science.

Beliefs about Science and Knowledge Integration

What is the relationship between beliefs about science and knowledge integration? Following our perspective on the development of scientific knowledge, it should be clear that students may hold a range of intuitions about the nature of science as well as about scientific phenomenon. As students develop their understanding of science and scientific phenomena, one goal is to integrate these views. In this study, we characterize consistencies between views of science and views of thermodynamics. In addition, we make conjectures concerning how these varied views might be integrated in productive and unproductive ways.

Students might integrate their views of science and of thermodynamics in an unproductive fashion. For example, if students believe that science is a collection of facts reported in textbooks, then memorizing those facts might be consistent with their view of science. In contrast, if students believe that science is progressing by fits and starts, and that scientists regularly reconsider ideas that have been developed in the past, then they might be inclined to understand scientific phenomena and to seek principles to explain their ideas. In addition, students who believe that science is dynamic might seek to integrate their diverse ideas and to build more predictive ideas about both science and thermodynamics.

Figure 1: Experiment Summary Card Indicating Principle and Prototype Integration



What we learned:

Our predictions were that ..the wool will do better

The reason we predicted this was .. because wool is thicker

But after this experiment we conclude that :the wool keeps the potato warmer then all the other conductors

Applying a principle:

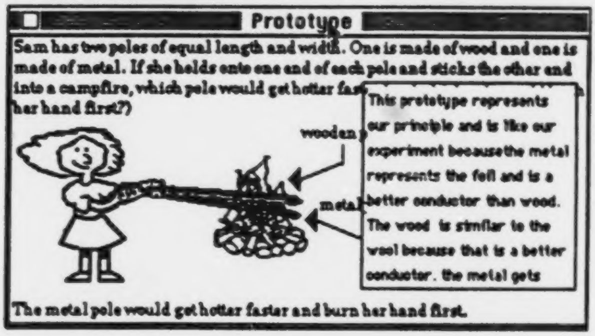
Our results are best explained by the following principles:

If two objects that differ only in material in which they are wrapped are placed in a cooler surround

(temperature) the temperature of the object with the better insulator around it will fall slower .

(heat energy) heat energy will flow out of the object with the better conductor around it faster .

When the Prototype button is pressed, a new window will open over the summary card with a prototype example such as the one shown below:



Sam has two poles of equal length and width. One is made of wood and one is made of metal. If she holds onto one end of each pole and sticks the other end into a campfire, which pole would get hotter fas
her hand first?)

This prototype represents our principle and is like our experiment becausethe metal represents the foil and is a better conductor than wood. The wood is similar to the wool because that is a better conductor. the metal gets

The metal pole would get hotter faster and burn her hand first.

Students' beliefs about science might empower them to participate in science courses and to integrate science knowledge. If students see scientists as grappling with complex ideas and trying to make sense of disparate information, they may feel empowered to engage in a similar process as they go about learning scientific ideas. In contrast, if they view scientists as quite different from themselves, they may decide to leave this grappling to scientists and wait until facts for them to memorize become available.

Views of science also have implications for application of scientific ideas to everyday life. If students view science as removed from themselves and their experiences, then they are unlikely to seek parallels between materials presented in their science classes

BEST COPY AVAILABLE

and their own experiences. This decision to separate knowledge acquired in science class from experience outside of class ultimately has disastrous consequences for the integration of scientific knowledge. Unless students also draw on their own experiences and information that they gather in other contexts, they are unlikely to build a realistic or predictive conception of science, and they are very likely to decide that science is irrelevant or at least unimportant to their own learning.

Summary

In summary, we conjecture that the development of scientific understanding proceeds from action knowledge, to intuitive beliefs, to scientific principles. A key feature of this development is the integration of ideas that may seem disparate or contradictory. In this article we characterize students' views of science as well as their views of thermodynamics, and examine the nature of the relationship that exists between these two forms of knowledge. We hypothesize that integrating views of science with knowledge of scientific phenomena will ultimately benefit students. For example, students who integrate their understanding of scientific phenomena across contexts such as school and everyday life may gain a more effective view of themselves as science learners and feel empowered to influence the direction science takes.

This study has three main goals. First, to characterize students' beliefs about the nature of science. Second, to assess the relationship between these beliefs and students' propensity to integrate knowledge of thermodynamics. Third, to examine the effectiveness of principles and prototypes as loci of knowledge integration and to determine which knowledge integration aid students prefer.

Methods

Subjects

The subjects of this study were 153 eighth graders enrolled in a one-semester physical science class in suburban northern California. The population was middle class, ethnically diverse, and contained a range of learning abilities. Students were assigned to one of six classes. All classes were taught by the same experienced mentor teacher.

CLP Curriculum

Students followed the Computer as Lab Partner (CLP) curriculum developed by a team of researchers including the classroom teacher. The design of CLP curriculum activities was guided by current research in cognitive science and science education. CLP activities emphasize hands-on experiments using computers which are on line to real-time data-collection devices such as temperature-sensitive probes (for more information, see Friedler, Nachmias, & Linn, 1990; Linn and Songer, in press b). Although the curriculum and cognitive goals have been reformulated and improved over the 11 semesters of its use, the science content, basic experiments, teacher role, and student characteristics have remained largely the same. In addition, a continuous objective has been to encourage the development of integrated understanding of the concepts heat energy and temperature.

The 12 weeks of instruction featured experiments investigating a variety of concepts including heating and cooling, insulation and conduction, temperature equilibrium, specific heat, relative rates of heat flow when volume, initial temperature or surface area change, and the general differences between the concepts heat energy and temperature. Students work in groups of 3 or 4 at one of 8 computers. For each experiment, the students determined the design, predicted the outcome, explained their predictions, conducted the experiment, reconciled the outcome with their predictions, and integrated the result. Recently, the curriculum activities have been expanded to include simulations written in HyperCard for the Macintosh. Students used the simulated experiments to gather data on activities that are difficult or dangerous to conduct in the classroom. All experimental data, both real time and simulated, is organized in an on-line laboratory notebook (Linn et al., in press).

Integration Interventions

Several activities were designed to encourage integration of thermodynamics concepts. These interventions fall into two categories: (a) the on-line integration interventions, which were identical for all students, and (b) the off-line integration interventions, which varied with class group. Integrations occurred at regular intervals during the semester.

On-Line Integration Interventions

At the conclusion of each experiment, students created an on-line summary card which included all of the important experimental components: their predictions prior to experimentation, the experimental features tested, and the results achieved, including written conclusions and their data graph. As part of the activities, students needed to use the experiment information contained on the summary card to both construct a scientific principle which summarized the experiment, and integrate the result with a prototype example of the same concept. To aid construction of the scientific principle, students were given prompts and feedback which were scaffolded in complexity—less general principles were accepted in early experiments, whereas only the most general principles were accepted in later experiments. Figure 1 displays a sample summary card containing a constructed principle and prototype. Therefore, in this semester, unlike previous semesters, all students integrated heat and temperature concepts through the construction of principles and the justification of prototypes as a regular and frequent part of their curricular activities. This integration practice came through participation in the on-line activities associated with each experiments' summary card.

Off-line integration interventions. In addition to the on-line integration activities mentioned above, students participated in two off-line integration interventions. These were: "Related Experiment Worksheets," and "Integration Homework."

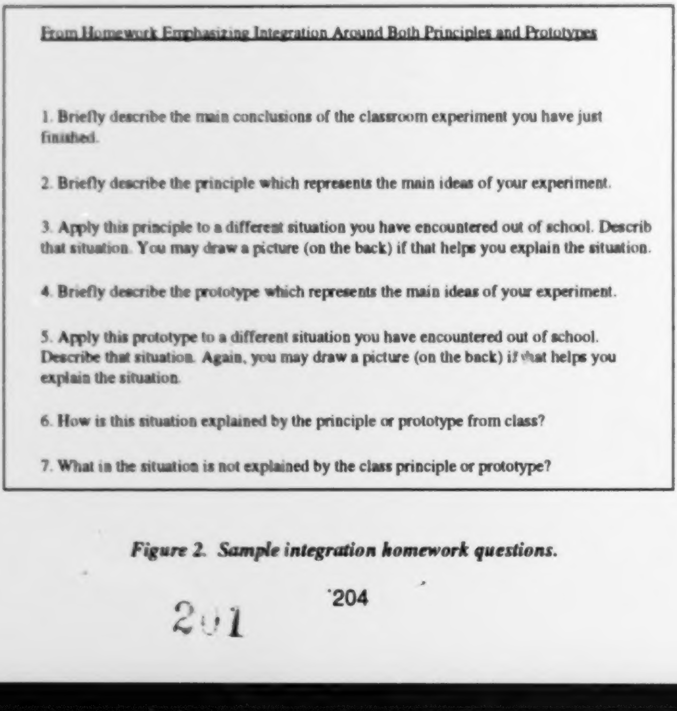
The Related Experiment Worksheets were used at four intervals throughout instruction, after each set of related experiments. For example, after all experiments concerning the effects of various insulating materials on the heating or cooling of liquids or solids, students were given an off-line worksheet which encouraged integration of the set of experiments. While all students were given instruction to integrate a set of experiments on each worksheet, the loci of integration emphasized varied by class: Two classes were instructed to integrate the experiments around principles, two around prototypes, and two around both principles and prototypes. Students were instructed to use the printouts of their experiment summary cards as background information.

The Integration Homework, given six times throughout the twelve weeks of instruction, required students to integrate one or two related experiments via the augmented loci of integration for their class (see Figure 2). For those classes emphasizing principles, students were asked to summarize the experiments, describe the principle constructed for that experiment set, and apply that principle to a new natural world situation which they felt was also explained by the class principle. Similarly, in those classes emphasizing prototypes, students summarized the experiments, described the prototype, and applied the prototype to a new situation. Students in the classes emphasizing both principles and prototypes for loci of integration did both of the activities as homework.

Experimental Design

All students participated in 12 weeks of experimentation using the CLP curriculum activities on the eight Macintosh computers in the classroom. Prior to instruction, all students were given two pretest evaluations: the Views of Science Evaluation and the Heat and Temperature Evaluation. During the instruction, students performed real-time and simulated experiments, summarized them through the use of summary cards, and worked on the off-line integration interventions involving worksheets and homework. In addition, a small number of students were chosen at random for elaborative interviews to assess subject matter knowledge and integration skills. After instruction, two posttest evaluations were given: The Heat and Temperature Evaluation and an additional Principle and Prototype Integration Evaluation.

The Views of Science Evaluation. The Views of Science Evaluation was developed to determine the character and stability of students' beliefs about the nature of science. The test consisted of 21 short-answer and true-false items in the following areas: students beliefs about the nature of science and scientific knowledge, beliefs about the role or work of scientists, and beliefs about what it means to learn science, both inside and outside of classroom situations.



From Homework Emphasizing Integration Around Both Principles and Prototypes

1. Briefly describe the main conclusions of the classroom experiment you have just finished.

2. Briefly describe the principle which represents the main ideas of your experiment.

3. Apply this principle to a different situation you have encountered out of school. Describ that situation. You may draw a picture (on the back) if that helps you explain the situation.

4. Briefly describe the prototype which represents the main ideas of your experiment.

5. Apply this prototype to a different situation you have encountered out of school. Describe that situation. Again, you may draw a picture (on the back) if that helps you explain the situation.

6. How is this situation explained by the principle or prototype from class?

7. What in the situation is not explained by the class principle or prototype?

Figure 2. Sample integration homework questions.

Designing items that elicit a range of valid responses has proven difficult. For several items responses varied little among all students, and for others, responses were not relevant to science beliefs. The nine items yielding valid, varied responses were used to determine student beliefs. Sample questions are included in Figure 3.

The data from the Views of Science Evaluation was used to create three science beliefs groups. Students who held the most productive beliefs about the nature of science were called the *dynamic beliefs* group. Of all the students, 15% were characterized as dynamic belief students. These students answered at least eight of the nine questions by indicating that they viewed science as understandable, interpretive, and integrated with many activities in the world around them. Those students who held the most unproductive beliefs were called the *static beliefs* group. We found that 21% of the students were characterized as static beliefs students. While all students expressed a few answers which implied productive beliefs, the static beliefs group demonstrated that for a majority of the nine questions analyzed, they largely viewed science knowledge as static, memorization intensive, and divorced from their everyday lives. The remaining group of students, called the *mixed beliefs* group, included the majority of students investigated. These students held some dynamic beliefs, some static beliefs, and some uninterpretable beliefs. Figure 3 demonstrates sample Views of Science Evaluation questions and characteristic responses for the static and dynamic beliefs students.

The Heat and Temperature Evaluation. The Heat and Temperature Evaluation was used to assess *isolated* thermodynamics knowledge and *integrated* thermodynamics knowledge.

Isolated thermodynamics knowledge questions assessed students' ability to interpret single experiments or natural world situations and make predictions. Many questions required students to provide examples of natural world situations, classroom activities, or scientific principles to support their answers. For example:

> Sam takes two identical pottery dishes filled with lasagna out of a hot oven and covers one with a pottery lid and the other with an identical metal lid.
>
> After five minutes, what temperature will each pottery dish have? (a) Both the same; (b) one with pottery lid warmer; (c) one with metal lid warmer.
>
> Choose yes or no for each of the following statements: (a) Heat energy will flow out of the dish with the pottery lid faster; (b) heat energy will flow out of the dish with the metal lid faster; (c) heat energy will flow out of both dishes at the same rate.
>
> What is the main reason for your answers?

Integrated thermodynamics knowledge questions assess students' abilities to synthesize the results from several experiments or activities. Students needed to go beyond the results of a given experiment or set of experiments in order to reach the higher levels of abstraction necessary for success with integrated thermodynamics questions. A sample question which emphasizes the distinction between the concepts heat energy and temperature is the following:

> In general, are heat energy and temperature the same or different? What is the main reason for your answer? Give an example which helps explain your answer.

Question	**Static Beliefs Students**	**Dynamic Beliefs Students**
When understanding new ideas, memorizing facts is better than trying to understand complicated material.	"Yes, when I was in 7th grade and we had an exam coming up, I would memorize facts and I would get a good grade on the test." "Yes, because if you try and understand complicated material there's a chance you won't understand it, with facts there's just facts."	"No, facts change." "No, sometimes the facts don't give you all the information you need."
Learning science for me is most like...	"Memorizing words and facts. That is how I learn science, that is how I learn it best."	"Doing puzzles, because in science sometimes the pieces don't fit in your head. Sometimes you don't understand it. Sometimes you do. You just have to keep on trying."
The science I learn in school has little or nothing in common with my life outside of school.	"Yes, when I drink orange juice, I don't break down the chemicals or minerals, I just drink it."	"No, sometimes I'll wonder why something in nature happened, and sometimes, with the little science I know, I can kind of make a guess."
Describe something you learned in science class which you could use to explain events outside of school.	"Nothing."	"I have learned how to read weather charts, how to use computers, different types of animals, and how to come up with more than one explanation for what happened." "If I found an animal or it's bones, I could probably identify it."
Describe something you learned in a science class that you will never use to explain events outside of school.	"Things about chemicals or animals."	"Nothing that I know of." "There isn't one. Everything you learn in science is based on true life."

Figure 3. Views of students holding static and dynamic beliefs.

The Principle and Prototype Integration Evaluation. The Principle and Prototype Evaluation assessed students' ability to integrate knowledge from classroom and natural world contexts around either scientific principles, prototypic examples, or both. Students were asked to apply and justify the use of scientific principles and prototypic examples to particular classroom and natural world situations. A sample principle and prototype question is as follows:

> You are preparing hot tea in a friend's kitchen. The kitchen sink contains two identical-looking faucets except one is made out of copper and the other is made out of a strong plastic of the same thickness. Your friend has the following question: When hot water is running out of both faucets and he wants to shut the water off, which faucet do you think will be LESS likely to burn your friend's hand?
>
> (a) Why do you think so?
>
> (b) Construct a principle which applies to this situation.
>
> (c) How is your principle similar to the situation?
>
> (d) Describe a prototype which applies to this situation.
>
> (e) How is your prototype similar to the situation?

This natural world situation resembled several classroom activities concerning insulation and conduction concepts that the students performed. In class, students conducted simulated experiments dealing with the rate of cooling of hot potatoes wrapped in materials with varying insulating abilities (aluminum foil, wool napkins, Saran wrap), the rate of heating Coke cans wrapped in similar materials, and a real time experiment where they tested the rate of cooling of hot water in glass beakers and Styrofoam cups. As a part of the concluding summary card activities, students were asked to construct a principle and justify a prototype. Similarly, the Principle and Prototype Integration Evaluation asked students to construct principles and describe prototype situations. In both the classroom and evaluation activities, students were judged successful at principle construction if they could design principles which identified the relationship between the changing variable and the rate of heat energy gain or loss. Students were successful with the prototypes if they could describe and defend a natural world situation which was different than the situation provided, but also identified the relationship between the changing variable and the change in heat energy.

Results and Discussion

To characterize students' beliefs about science, we separated students into the static, mixed, and dynamic groups based on their responses to the Views of Science Evaluation. We then examined the performance of these groups on the various knowledge integration assessments. In addition, we examined whether the knowledge integration augmentations improved understanding. We also assessed students' preferences for prototypes or principles as the locus of knowledge integration.

Views of Science

Only about 15% of these middle school students reported a dynamic view of science. In contrast, 21% reported static views, and the remaining 63% held mixed views. As shown in Figure 4, our criteria for static and dynamic views of science were relatively

204

strict. Only students who consistently answered questions from a particular perspective were assigned to these groups. Most students ended up in the mixed beliefs group.

Students with dynamic views of science recognize that scientific knowledge is controversial. These students realize that scientists compare results and that scientists can look at the same experiment and reach different conclusions. Students in this group also believe that scientists use evidence to resolve controversy. As would be expected, students in this group extend their view of the dynamic nature of science to their perspective on science learning, reporting that science principles in textbooks may not be true, and that understanding new ideas is preferable to memorizing facts when trying to learn complicated material. These students have a cohesive and realistic view of science. They are likely to understand instruction that stresses that science is socially constructed. We anticipate that, in addition, these students recognize the possibility that if they were to participate in science, they could have an impact on the direction of scientific advance.

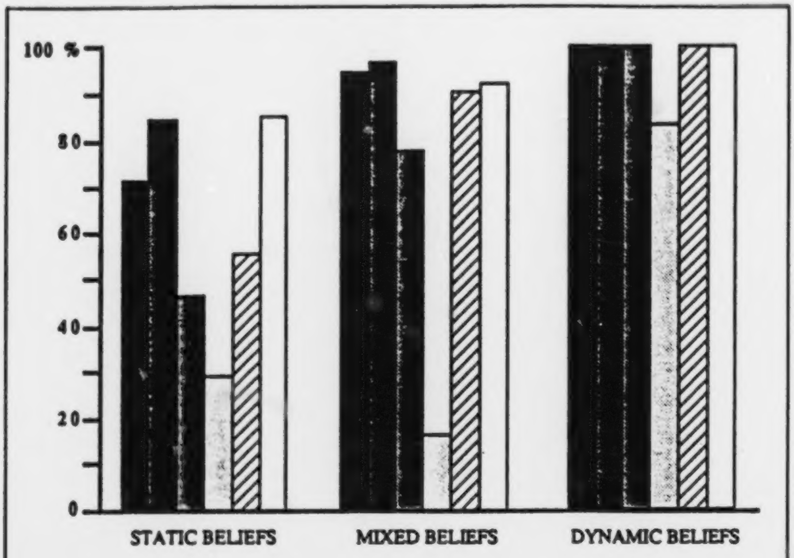
In contrast, the 21% of students who hold static beliefs about the nature of science believe that scientists do not expect principles to explain a broad array of events. These students deny the integrative function of scientific knowledge acquisition. Furthermore, they cannot differentiate between established scientific and current scientific controversies. In addition, these students often believe that all scientific principles in textbooks will always be true, and they view science as best learned by memorizing facts rather than attempting to understand complicated material. Thus, these students see science as essentially static and unchanging. They see scientists as simply adding to the store of knowledge rather than debating alternative perspectives or trying to group many events under a single principle.

How do these belief groups reflect our perspective on the development of scientific knowledge? The small number of students who hold dynamic beliefs about the nature of science have an intuitive view of science that is consistent with that of experts. These students have integrated their view of science with their ideas about learning science. Furthermore, their perspectives of scientists are consistent across questions about scientific knowledge and scientific behavior. These students recognize the integrative function of scientific principles.

In contrast, students with static beliefs about the nature of science appear to hold quite a few beliefs that are at the level of action knowledge. We suspect that these beliefs stem from experience with classroom science learning. These students' knowledge reflects primarily a rather superficial perspective on science learning. These views lack integration with students' everyday experiences. For example, students' static beliefs are incompatible with the notion that scientists at one point believed that the world was flat, or with debates among scientists concerning the explanations for earthquakes or the extinction of dinosaurs.

Students with mixed beliefs have a variety of ideas about the nature of science and about the learning of science. This is consistent with our findings of mixed beliefs in other domains. Our examination of students' views of thermodynamics, for example, reveals that students often hold unintegrated views of heating and cooling. Thus, students may assert that wool is a good material for keeping people warm, but is not able to keep ice cream cold. Students may draw on their experiences outside of science class to justify some of their intuitions, while drawing on explanations from science

Figure 4: Responses to Beliefs Quiz Questions by Beliefs Groups



Students chose Yes or No in response to the following statements:

	Beliefs about scientists:	Recorded Answer
■	Scientists can look at the same experiment and reach different conclusions.	Yes
■	Scientists expect one principle to explain many scientific events.	Yes
■	Scientists disagree about explanations for scientific events.	Yes
░	To verify their findings, scientists compare their results to those of others.	Yes
	Beliefs about learning science:	
▨	The science principles in the textbooks will always be true.	No
☐	When understanding new ideas, memorizing facts is better than trying to understand complicated material.	No

textbooks that report the results of the scientific process rather than describing the nature of scientific events to justify other intuitions.

In summary, most students have unintegrated views about the nature of science. Some of those views may stem from a superficial interpretation of their science textbooks. Students who view the information in a science textbook as static are unlikely to examine the process that generated the information. Students with dynamic views of science draw on their understanding of the process of scientific advance. These students generally hold more productive views about the learning of science than students with static views of science. Should science educators help students develop a more dynamic view of science? Would such a view facilitate learning of scientific concepts? To partially answer this question we examine the relationship between beliefs about science and knowledge integration.

Knowledge Integration

To investigate the relationship between views of science and the integration of thermodynamics knowledge, we summarized performance on the knowledge integration assessments and compared the performance of the three beliefs groups. As discussed in the methods section, instruction featured the CLP curriculum reformulated to emphasize knowledge integration. Students were encouraged to integrate their knowledge around pragmatic principles and prototypic events during on-line integration activities associated with each experiment. All students used the pragmatic principles and the prototypic events to summarize the experiments that they conducted in class (see Figure 1). In addition, during off-line integration activities and homework assignments, some students integrated their knowledge around principles, others used prototypes, and a third group used both principles and prototypes. In this section, we look at knowledge integration independent of the off-line integration augmentations. The following section will examine the effect of varying integration augmentations on student knowledge integration.

We examined three aspects of learning from the Computer as Lab Partner curriculum. First, we looked at students' understanding of thermodynamics ideas in isolation from each other. Second, we looked at two forms of knowledge integration: first, integration of information around principles and prototypes, and second, integration of heat energy and temperature.

All students made substantial gains in their understanding of general concepts in thermodynamics as shown in Table 1. In particular, after instruction, 89% of the students were successful in predicting the outcome of an everyday situation similar to an experiment conducted in class. Also, 77% of the students were successful at explaining the reasoning behind their prediction.

Overall, 60% of the students successfully used principles and prototypes introduced in the study of thermodynamics to integrate their knowledge from different contexts (Table 1). When integration of information around principles and prototypes was analyzed by students in beliefs groups, students in the static beliefs group were somewhat less successful than students in the dynamic beliefs group. Comparing static beliefs students to dynamic beliefs students for the five principle/prototype integration items, the effect size was 0.45, which shows a trend towards greater integration by dynamic beliefs students; however this result was not statistically significant ($df = 2$, $n = 137$, $p < 0.21$).

Finally, all students made substantial gains in their understanding of the distinction between heat energy and temperature. The pretest-posttest effect size was 0.4 (Wilcoxson Signed-Ranks Test, $p < 0.05$). Students on the posttest ($m = 0.31$, $sd = 0.46$) were significantly more successful at heat and temperature integration than students on the pretest ($m = 0.16$, $sd = 0.37$). In addition, on the heat energy and temperature interpretation item, those with dynamic beliefs were significantly more successful ($m = 0.59$, $sd = 0.50$) than those with static beliefs ($m = 0.13$, $sd = 0.35$). The effect size was 1.09 (see Figure 5). Furthermore, on the five integration items combined, the same relationship held with dynamic ($m = 4.15$, $sd = 1.05$) exceeding static ($m = 3.03$, $sd = 1.45$) by an effect size of 0.92. An ANOVA comparing the three belief groups yielded a significant effect as well (df = 2.143, $p < 0.01$).

Thus, as anticipated, there is a relationship between beliefs about the nature of science and students' propensity to integrate the knowledge presented in the CLP curriculum. Students with dynamic beliefs about science were more likely to develop integrated understanding of heat energy and temperature than students with static beliefs.

For students with dynamic beliefs about the nature of science, we see an integration and cohesiveness between their views of science learning and the kinds of material that they master. These students believe that the best way to learn science is to understand it and they recognize that scientists expect one principle to explain many scientific events. Their understanding of science is consistent with this perspective. They are more likely to have integrated views of heat energy and temperature, and to attempt to organize the material presented to them around principles and prototypes.

Consistency with Other Investigations

This finding extends the work of Songer (1989). Songer found that students with predictive beliefs about the nature of science performed better on integration than those with unpredictive beliefs. She also reports that type of belief was not associated with accuracy of recall of heat and temperature information.

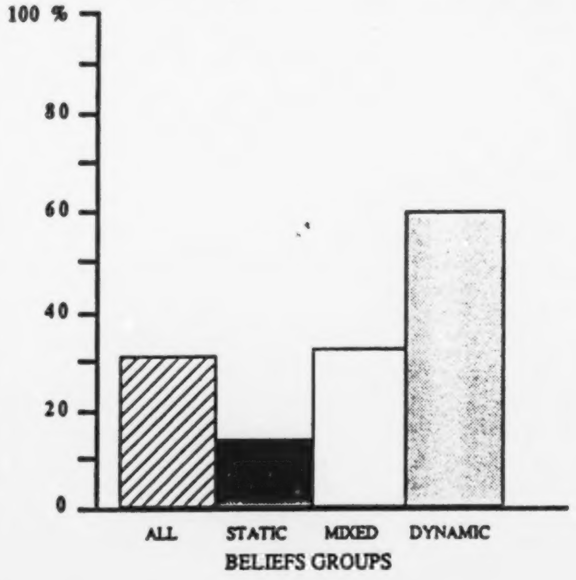
Thus, in both the Songer investigation and in this investigation there was no relationship between beliefs about science and recall of heat and temperature information. These findings are consistent with the view that students who hold static views of science and memorize information will do just as well on tests that do not require knowledge integration as will students who are attempting to develop integrated understanding. In contrast, when integrated understanding is emphasized in the curriculum and required on the assessments, then students with dynamic views of science will be more successful than students with static views.

Table 1

Progress on Knowlege Integration

	Correct responses (n = 153)	
Question type	Percent	Standard deviation
Predict outcome for everyday situation	89.0	0.31
Explain prediction for everyday situation	77.6	0.42
Construct a principle to explain class of everyday situations	57.7	0.50
Apply a principle to a new situation	59.6	0.50
Create a prototype to explain class of new situations	50.9	0.50
Apply a prototype to a new situation	60.2	0.49

Figure 5: Performance on Heat and Temperature Integration Question by Beliefs Groups



Question: Explain why heat energy and temperature are the same or different. Give an example which helps illustrate your answer.

(one way ANOVA, $p \leq .002$)

Locus of Integration: Principles and Prototypes

A third goal of this investigation was to examine the effectiveness varying the locus of integration for related experiment worksheets and integration homework. After every two-three experiments, students completed the off-line integration homework activities that emphasized either principles, prototypes, or both, as the locus of integration. In addition, students practiced organizing these related experiments around principles, prototypes, or both in off-line related experiment worksheets. Two classes were encouraged to integrate their experience around principles, two classes focused on prototypes, and two classes used both principles and prototypes.

These integration activities were very successful regardless of which loci of integration was emphasized. Students completed the assignments and reported enjoying the activities. In general, students liked all of the approaches. There were no differences between students assigned to the principles group, the prototypes group, or the principles-plus-prototypes group on subject-matter knowledge acquisition or on the ability to integrate information about thermodynamics. This finding is not really very surprising given the limited nature of the off-line activities as well as the fact that all students were exposed to principles and prototypes during regular class activities.

We had conjectured that prototypes would be more appropriate for students whose knowledge of thermodynamics was primarily at the action level, whereas principles would be more suited to students who had intuitive beliefs about thermodynamics. To gain insight into this conjecture, we asked students to indicate which of the integration loci they preferred. The results of this survey are reported in Figure 6.

Students, in general, preferred principles over prototypes. However, students whose off-line integration activities primarily involved prototypes were far more enthusiastic about prototypes than were other students. Essentially, when asked whether they would eliminate principles or prototypes from future class activities, 68% of the students who used principles to integrate their experiences would eliminate prototypes, whereas 66% of the students who used prototypes to integrate their experiences would eliminate principles. To clarify these results, we interviewed students and asked them why they preferred principles or prototypes for knowledge integration. Students favoring prototypes frequently indicated that the vocabulary in the principles was too abstract and difficult, whereas students favoring principles over prototypes frequently stressed that the principles were more general and useful.

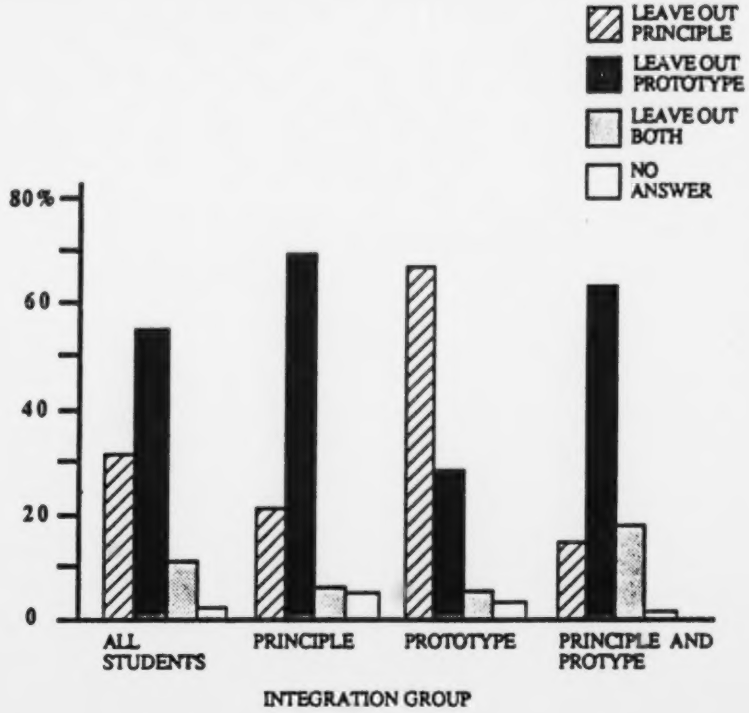
When asked to explain why they recommended leaving principles out of the curriculum in the future, many students focused on the concreteness of prototypes. Examples include:

> Student (S1): I think the prototype is a little better.
>
> Interviewer (I): Why do you think so?
>
> S1: It is something you can relate to.
>
> I: What do you mean by 'relate to'?
>
> S1: Like, there is different ones where they put something on fire, wood, and metal and you can understand why you would heat up certain things from past experiences.
>
> I: So you kind of know something like that from outside of science class?
>
> S1: Yes.
>
> I: Okay. So, how was that different from what the principles do?
>
> S1: Well, some people might not understand...it just has some words that some people might not understand the words but they may understand the idea of it.
>
> I: So there is some words in there that were kind of confusing or what do you mean by they understood the idea?
>
> S1: They could understand the concept of it but not how to word it.

Other students preferred prototypes because they found the concrete examples particularly illuminating. They responded, for example:

> I: ...Can you think about which one [principle or prototype] you know, if you had plenty of time, you might like to keep?
>
> S2: Prototype, because it's fun.

Figure 6: Preferences for Principles and Prototypes by Knowledge Integration Group



Question: If for next year's class we needed to leave out some parts of the materials used in this class, which one of those listed would you suggest we drop?

(one way ANOVA, $p \leq .0001$)

I: Okay, good. So it's more fun. How about what helps you learn more? Either of them help you understand the experiment any more?

S2: Prototype does.

I: Okay. What is it about the prototype that helps you understand more?

S2: When you're creating something [constructing the principle] it's kind of hard to get the actual process of how it works, but if you're actually doing it, it's a lot easier, you can see how it's been done and why.

I: So you mean the prototype is talking about something you can actually do?

S2: Uh huh. You know, you can be there to see what happened.

(later in same interview)

> I: ...Okay, So then, if this [prototype] is the whole thing, what is missing from the principle?
>
> S2: Um...probably just being right there and seeing this one in depth. You can be right there. That [the principle] just tells you the main facts.
>
> I: Okay, so the principle is just talking about facts, whereas the prototype gives you what?
>
> S2: Everything.

Students who preferred principles over prototypes emphasized that the principles were clear and concise statements with broad explanatory power saying, for example:

> S3: I think you should drop the prototype because most of the time, you just finished the lab, you know pretty much about it, except the principle, you have to construct then, you have to figure out which one is right and which one is wrong and how, and once you get that, then you have plan to a prototype which you already know, which is irrelevant because reiterating what you learn.
>
> I: It's not giving you any new information?
>
> S3: No, not really.
>
> I: What does the principle do?
>
> S3: It makes you think, in what situations it will work, well, a conductor will rise or fall more rapidly, but if you just put down fall, what about rise, it makes you think more, more universal. Prototypes just makes you think of instances and what is similar.
>
> I: Can you say more about what you mean when you say that it [a principle] makes you think more?
>
> S3: When you first decide what to write down or choose for the principle, there is so many choices and you have to think of what is universal because you think of, if you think that heat will fall, then you will think that it will fall faster or slower, then it says it will rise or fall then you think, well, maybe we'll do the same when it starts cold and goes to room temp. So maybe you think, "Hey, maybe rise or fall in any condition..." and you just think about other variables besides the one you are working on.

Others who preferred principles talked about the usefulness of principles for making predictions:

> S4: And a principle, I think is more important.
>
> I: Why?
>
> S4: And also principle refers to everything, prototype also goes under principle.
>
> I: So you can use the principles to talk about the prototypes too?
>
> S4: Yes.
>
> I: How does that make principles better?
>
> S4: When I do experiments I look at the principle a lot because I think it helps me understand experiments more and how to predict and what it's all about and what we are trying to find. If I look at prototype, it explains the situation and it

> explains what we are trying to look and what is wrong but it doesn't tell me exactly why or I don't know how to explain it. But the principle just tells you what it's all about and it explains everything.

These results suggest that whereas most students prefer principles for integrating their knowledge, there are students who find prototypes to be preferable. Furthermore, students may need considerable experience using prototypes before they fully recognize the advantages of these prototypes. Only those students who had augmented experience using prototypes to integrate their understanding recognize the value of prototypes for their own learning.

These findings provide insight both into the effectiveness of knowledge integration aids and students' knowledge of their own integration process. As can be seen, many students indicate they prefer principles over prototypes for integrating their knowledge; however, they may learn more by using prototypes than principles. At least for students in the prototype off-line integration condition, there was a gain in the realization of the importance of prototypes for knowledge integration. Instruction needs to help students both understand prototypes and recognize how to use prototypes to organize their knowledge. Students whose understanding of scientific phenomena is quite concrete may need help in integrating prototypes with their concrete knowledge in order to reach broader views of the scientific domain.

These results also demonstrate the advantages of using alternative integration aids for increasing student understanding of scientific phenomena. Some students preferred principles, whereas others found prototypes most effective. Overall, it was difficult to distinguish the effects of principles from the effects of prototypes. However, for individual students it was quite apparent that principles helped some, while prototypes helped others. The constructive process is an idiosyncratic one. Providing students with options within the learning environment is clearly advantageous and most likely to foster construction of understanding across a broad range of learners in complex domain such as thermodynamics.

Conclusions and Implications

These results strongly suggest that students' beliefs about science are important and complement the development of thermodynamics knowledge. These studies also underscore the importance of focusing specifically on knowledge integration in helping students understand scientific phenomena. Students with predictive beliefs about science develop integrated knowledge. At the same time the process of integrating knowledge illustrates the complex process of scientific advance. Students may not integrate material presented in science classes if they believe that science consists of isolated principles. In addition, students may not develop a view of science consistent with historical evidence if science is presented to them as a collection of fairly unrelated facts and ideas. Thus, students are likely to gain a more integrated understanding of science if their courses emphasize both knowledge integration and knowledge about the nature of science. Ultimately the goal is to help students integrate their beliefs about science with their understanding of scientific phenomena. Only with this view can students appreciate that the development of scientific knowledge results from the activities of those in science.

This study demonstrates the advantage of explicit emphasis on integrated understanding. Songer (1989) demonstrates the effectiveness of the continuum line for helping students to integrate their understanding of insulation and conduction. This study augments and extends that investigation to demonstrate that principles and prototypes are suitable loci for the integration of groups of experiments in thermodynamics. This study also suggests that offering students a range of loci for integrating their understanding is advantageous, because for some students, principles would be too abstract, and for others, prototypes would be too narrow.

Goals for Science Instruction

This study illustrates the danger of focusing science instruction too narrowly on facts or isolated pieces of scientific knowledge. Students rarely spontaneously integrate information presented in isolation. Instruction needs to focus students on constructing integrated understanding and support them in the process of developing these integrations. In order for students to move beyond isolated ideas and into a more predictive and productive understanding of science, intervention is needed. Early versions of the computer as lab partner curriculum did not foster sufficient knowledge integration and students were unable to organize their knowledge around principles and prototypes (Friedler et al., 1990; Linn & Songer, in press b). Instruction specifically focusing on integration was successful. This instruction supported students in their efforts to construct robust and cohesive views of thermodynamics. In addition, analysis of student learning under these conditions illustrated that dynamic beliefs about science were associated with more integrated understanding. Presumably, further augmentation of the curriculum to specifically emphasize the integration between views of science and views of thermodynamics would have even greater benefits for students. Such approaches are described by Duschl and Gitomer (1991).

Integration with Other Investigations

These investigations coincide with findings from other researchers and serve to augment and integrate previous investigations. Many researchers report that students have unproductive beliefs of science. Schoenfeld (1983) notes that students learning mathematics often report that the main idea is to memorize algorithms and procedures rather than to try to make sense of the problems or the results of their problem-solving efforts. Schoenfeld also reports that students with more productive views of mathematics are more likely to make sense of mathematical problems and to report reasonable solutions. Students who lack this view of mathematics often give answers to problems that make no sense, suggesting for example that to transport a group of individuals you would need "three buses remainder six" or that the amount of wood needed to satisfy a set of constraints is "minus six" feet. Schoenfeld calls on mathematics educators to help students make sense of mathematics learning rather than allowing them to memorize mindless algorithms.

Novak and his colleagues have focused on knowledge integration from the perspective of having students draw concept maps. These investigators report that students often represent their knowledge as isolated and incohesive. Novak calls on teachers to review these maps and to work with students to help them recognize how information can be integrated (Novak, Gowin, & Johannsen, 1983). Schommer (1990) reports similar relationships by assessing students' beliefs about the nature of reading and their performance on reading comprehension tasks. Students who believe that comprehension is at the level of words or sentences are less likely to make sense of written communication.

Taken together, these investigations suggest the importance of helping students (and teachers) understand the processes under which knowledge is generated as well as the results of those processes. Unless students look jointly at knowledge generation and the outcomes of the knowledge generation process, they are unlikely to develop an integrated, productive understanding of science or of science concepts. Thus we call on educators, curriculum designers, researchers, and textbook writers to design materials that help students construct integrated understanding in a broad range of domains. We note, for example, that the Harvard Project physics curriculum emphasized a historical perspective on knowledge generation. Students responded favorably to this approach, yet it has not received widespread acceptance in textbooks and is rarely found in materials for mathematics and other domains.

Often we hear that historical examples about the process of scientific advance are left out of science courses because there simply is not enough "instructional time" to cover everything (Eylon & Linn, 1988; Linn, 1987). We believe that this perspective is wrong headed. Morrison (Apelman, Hawkins, & Morrison, 1985) has argued that "less is more." Unless students have sufficient opportunity to understand the nature of the knowledge generation process, they are unlikely to become participants in this process in the future and may instead believe that science knowledge is irrelevant to their own lives and to the lives of others like them.

Acknowledgments

Authors' Note: This material is based upon research supported by the National Science Foundation under Grant No. MDR-88-50552. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors appreciate helpful discussions with other members of the Computer as Lab Partner Project. Special thanks go to project member Doug Kirkpatrick, who assisted in the design of the curriculum and taught the classes, to Judy Stern, who coordinated the development of the software, and to Eileen Lewis, whose longitudinal investigations complement the reported findings. Thanks also to Evelyn Shapiro and Darragh Perrow for editing and producing this document.

References

Apelman, M., Hawkins, D., & Morrison, P. (1985). *Critical barriers phenomenon in elementary science*. Grand Forks, ND: University of North Dakota, Center for Teaching and Learning.

Brecht, B. (1980). *The Life of Galileo*. London: Eyre Methuen.

Clement, J., Brown, D., & Zietsman, A. (1989). Not all preconceptions are misconceptions: Finding "anchoring conceptions" for grounding instruction on students' intuitions. *International Journal of Science Education, 11*, 554-565.

Duschl, R.A., & Gitomer, D.H. (1991). Epistemological perspectives on conceptual change: Implications for educational practice. *Journal of Research in Science Teaching, 28*, 839-858.

Eylon, B., & Linn, M.C. (1988). Learning and instruction: An examination of four research perspectives in science education. *Review of Educational Research, 58* (3), 251-301.

Friedler, Y., Nachmias, R., & Linn, M.C. (1990). Learning scientific reasoning skills in microcomputer-based laboratories. *Journal of Research in Science Teaching, 27* (2), 173-191.

Grosslight, L., Unger, C., & Jay, E. (1991). Understanding models and their use in science:

Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching, 28*, 799-822.

Kuhn, T.S. (1970). *The structure of scientific revolutions* (2nd edition). Chicago: University of Chicago Press.

Linn, M.C. (1987). Establishing a research base for science education: Challenges, trends, and recommendations. *Journal of Research in Science Teaching, 24* (5), 191-216.

Linn, M.C., & Songer, N.B. (in press a). Cognitive and conceptual change in adolescence. *American Journal of Education* (Special issue: Development and Education Across Adolescence).

Linn, M.C., Songer, N.B., Lewis, E.L., & Stern, J. (in press). Using technology to teach thermodynamics: Achieving integrated understanding. In D.L. Ferguson (Ed.), *Advanced technologies in the teaching of mathematics and science*. Berlin: Springer-Verlag.

Metz, K.E. (1991). Development of explanation: Incremental and fundamental change in children's physics knowledge. *Journal of Research in Science Teaching, 28*, 785-797.

Minstrell, J. (1982). Explaining the 'at rest' condition of an object. *The Physics Teacher, 20*, 10-14.

Novak, J.D., Gowin, D.B., & Johannsen, G.T. (1983). The use of concept mapping with junior high school science students. *Science Education, 67*, 625-645.

Piaget, J. (1952). *The origins of intelligence in children*. New York: National Universities Press.

Reif, F., & Larkin, J.H. (1991). Cognition in scientific and everyday domains: Comparison and learning implications. *Journal of Research in Science Teaching, 28*, 733-760.

Schoenfeld, A.H. (1983). Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance. *Cognitive Science, 7* (4), 329-363.

Schommer, M. (1990). The effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82* (3), 498-504.

Songer, N.B. (1989). Promoting integration of instructed and natural world of knowledge in thermodynamics. Unpublished doctoral dissertation, University of California, Berkeley, CA.

STRATEGIC TEACHING IN SCIENCE

Charles W. Anderson

This chapter discusses cognitive instruction in science; that is, teaching science in a way that helps students monitor and improve their thinking about the natural world. Much of this chapter is based on a program of research and development in which I and a number of colleagues (especially Edward L. Smith and Kathleen J. Roth) have been engaged for the last seven years. In this program, we have studied science teaching and learning at the elementary school, middle school, and college levels, and we have developed and implemented new approaches to teaching science at each level.

Our program is but a small part of a substantial research tradition involving researchers in many countries who have studied the teaching and learning of science from a cognitive perspective. This research tradition has produced books such as those by Driver, Guesne, and Tiberghien (1985) and by West and Pines (1985), as well as a variety of review articles such as those by Anderson and Smith (1987) and by Carey (1986).

In writing this chapter I was asked to address three questions:

1. To what extent is the framework for strategic teaching and learning that was developed in the first half of this book consistent with research in science?

2. To what extent is this framework helpful for planning sequences of instruction in my content area?

3. What adaptations, if any, need to be made to teach low- and high-achieving students so that all students will benefit from strategic teaching and cognitive instruction?

The chapter ends with direct answers to these questions. My answers are based on my understanding of the problems science teachers encounter in planning and in the classroom. Therefore the chapter begins with a discussion of those problems and of some of the solutions that we and others have developed. This discussion focuses on the problems encountered by a hypothetical teacher whom I shall call Ms. Lane. She is a fifth-grade teacher who is preparing to teach a unit on light and vision. Although she

---

This selection first appeared as a chapter in B.F. Jones, A.S. Palinscar, D.S. Ogle, and E.G. Carr, Eds. (1987), ***Strategic Teaching and Learning: Cognitive Instruction in the Content Areas****, pp. 73-91. Alexandria, VA: Association for Supervision and Curriculum Development. Copyright © 1987, North Central Regional Educational Laboratory (NCREL), Oak Brook, IL. Reprinted with permission.*

is not familiar with the term "strategic learning," she shares many of its goals. In particular, she is concerned about her students' *understanding* of the content she teaches. She remembers science courses that she survived by memorizing for tests, then soon forgetting what she had memorized. Too often, she sees her students starting to do the same thing; they seem to come out of science lessons with memorized trivia rather than true understanding, and this is a matter of great concern to her.

Ms. Lane is dissatisfied when she sees her students memorizing rather than understanding, but what can she do to help them? This clearly is not an easy task; if it were, there would be a lot less memorizing in school science classes. In order to teach for understanding, Ms. Lane must develop solutions to a number of difficult problems. The following sections focus on four of those problems:

1. A definitional problem: What does it mean to "teach for understanding" in science?
2. A curricular problem: What is it about science that is worth understanding
3. A student learning problem: How does scientific understanding develop in students?
4. An instructional problem: What can teachers do to help students understand?

The Definitional Problem: Teaching for Understanding

Teaching for understanding is, of course, what this book is all about. Strategic teaching assumes that what teachers teach will be based in part on what they understand about how students learn. This focus on understanding is apparent in the six principles about learning with understanding that form the framework for Chapter 1 of this book. Restating these principles may aid science teachers like Ms. Lane to understand what it means to teach for understanding in science.

1. Learning is goal oriented; skilled learners are actively involved in constructing meaning and becoming independent learners.
2. Learning is linking new information to prior knowledge.
3. Learning is organizing knowledge.
4. Learning is strategic; skilled learners must develop a repertoire of effective learning strategies as well as be aware of and control their own activities.
5. Learning occurs in phases, yet is nonlinear and recursive; learners must think about what they already know, anticipate what they are to learn, assimilate new knowledge, and consolidate that knowledge.
6. Learning is influenced by development, and there are important developmental differences among learners.

Ms. Lane is not an expert on educational research. She is, however, an experienced teacher who has developed her own ideas about understanding through years of practice. She is a good teacher, and as she compares her ideas with those listed above, she finds that they are highly compatible. (Unfortunately our research indicates that this is not universally true. Hollon and Anderson (1986), for example, found that some experienced teachers, typically teachers who do not understand science very well themselves, found the ideas above threatening and incompatible with their present practice.)

Therefore, Ms. Lane sets a goal for herself. She accepts the ideas listed above as a reasonable definition of learning with understanding, and she resolves to teach in a way that helps her students show these characteristics in their own learning. Before she can achieve this goal, however, she faces important problems involving curriculum, student learning, and instruction.

The Curricular Problem: What Is Worth Knowing?

At the outset of curriculum planning, Ms. Lane and other science teachers must decide what they want their students to learn. It seems reasonable to expect that textbooks are not as helpful in this respect as they should be. When Ms. Lane opens her textbook to the unit on light, she discovers that the textbook presents only a deceptive appearance of having clearly defined "content" that can be taught to students in a straightforward manner. In fact, she can interpret the textbook from a variety of different perspectives, and each perspective suggests a different definition of the curricular content of the light unit. Three perspectives that are common among teachers and science educators are contrasted below.

Teaching Science as Facts, Rules, and Definitions

When she looks at her textbook, the first thing that Ms. Lane notices is that it contains a lot of facts (light travels in straight lines; light travels very fast; the colors in the spectrum...). There are also some definitions (transparent, translucent, opaque, lens, retina...). Although this textbook does not have them, it is also common for more advanced texts to contain formulas and rules. It certainly seems reasonable, looking at a science textbook, to think of science as a long list of facts, rules, and definitions to be learned. The most reasonable interpretation of Ms. Lane's textbook is certainly that the authors intended science to be taught this way.

Writing textbooks that present lots of facts makes good business sense. This way of thinking about science is common among the teachers we have worked with in our research and development projects (Anderson and Smith in press, Hollon and Anderson 1986). Even college professors often treat science teaching essentially as the presentation of facts, rules, and definitions, albeit in a well-structured and organized manner (Anderson 1986). Thus teaching science by presenting facts, rules, and definitions represents the "path of least resistance" for both teachers and textbook publishers.

Unfortunately, it is virtually impossible to teach lists of facts in the meaningful way that we defined as cognitive instruction at the beginning of this chapter. Learning facts puts the students in a passive rather than an active role, encourages memorization rather than the active construction of knowledge, and fails to connect science content with students' prior knowledge of the world. Because of these deficiencies, attempts to teach science as facts, rules, and definitions typically produce little understanding and less retention (cf. Anderson, Sheldon, and DuBay forthcoming; Anderson and Smith in press; Clement 1982). Ms. Lane knows from experience that she needs to do more than teach her students lists of facts, rules, and definitions. There *must* be more to science than that, but what?

Teaching Science as Process Skills

One popular answer to Ms. Lane's question focuses on "science process skills." These are the skills necessary to conduct investigations of the natural world, such as observing, measuring, inferring, and so forth. When Ms. Lane looks at her textbook, she sees that in addition to facts and definitions it has "finding out" activities. Although the activities in this book are not particularly good, it would be possible to modify these experiences and use them to provide students with hands-on laboratory activities that would provide training and practice in science process skills. Would this be enough to help Ms. Lane's students learn science with understanding?

Unfortunately, the research evidence suggests that hands-on activities or instruction in process skills will not ensure meaningful learning, either alone or in combination with conventional fact-based instruction (Roth 1984, Smith and Anderson 1984). One problem is that science processes do not seem to consist of unitary skills that can be transferred from one context to another. Observing cell cultures, for example, has little in common with observing geological formations or with observing chemical reactions. Furthermore, a major component of process skills seems to be content knowledge (e.g., a good observer of cell cultures must know a lot about cells).

So science is more that the sum of process skills plus facts, rules, and definitions. Ms. Lane must find a different way of thinking about the content that she teaches before she can truly engage in cognitive instruction in science.

Teaching Science as Explanation

There is another approach to defining the content of science, one that has much more potential for serving as the basis for cognitive instruction in science. This approach emphasizes the explanatory functions of science. Science is our best attempt to explain *how* and *why* events happen as they do in the natural world. How does light help us to see, for example? How do we see colors? Why do plants need light to grow? Why is the sky blue?

If Ms. Lane looks hard, she can find passages in her textbook in which scientific principles are used to explain natural phenomena. Here, for instance, is a passage explaining how light helps us to see:

> Like a rubber ball, light bounces off most things it hits. When light travels to something opaque, all the light does not stop. Some of this light bounces off. When light travels to something translucent or transparent, all the light does not pass through. Some of this light bounces off. When light bounces off things and travels to your eyes, you are able to see (Blecha, Gega, and Green 1979, p. 154).

Is this a "main point" of the unit on light? There is certainly no indication that the authors of the text thought of it that way. The passage quoted above is buried in the middle of a chapter; no special markers indicate its importance. Our research, though, has convinced us that Ms. Lane and other science teachers can truly teach for understanding only if they "reconstruct" the contents of science texts, putting much more emphasis than most texts do on using scientific theories and principles to explain phenomena in the natural world.

There are a number of advantages to focusing on science as explanation. This approach gives back to scientific theories their original purpose, a purpose that all too often is lost in science as it is taught in schools. Scientists develop their theories because they have *questions* about how and why the world works. All too often, we teach students the answers (or procedures for getting answers) without letting them know what the questions are. Recent work in the history and philosophy of science has emphasized that science is fundamentally an attempt to describe and explain the natural world. Scientific progress involves constructing more powerful descriptions and explorations, not simply discovering more facts and laws according to the "scientific method" (cf. Kuhn 1970, Mayr 1982, Toulmin 1961).

So Ms. Lane decides that she will emphasize the explanatory functions of science in her treatment of the light unit. She will ask questions calling for explanations and help her students use passages from the text like the one above to construct scientific answers to those questions. In making this decision, Ms. Lane has taken an important step toward cognitive instruction in science. She has decided to involve her students in a central activity of science, one that requires them to actively construct meaning and to organize and use their personal and scientific knowledge. She has also taken a key step toward linking the information in the light unit to her students' prior knowledge. This linkage is examined in detail in the next section.

The Student Learning Problem: How Does Scientific Understanding Develop?

Ms. Lane has now made a strategic curricular decision about her unit on light and vision: she will emphasize teaching her students to explain phenomena scientifically. As she looks through the textbook, she can now see a number of phenomena that she wants her students to explain: how we see; how light interacts with transparent, translucent, and opaque objects; how our eyes work; and so forth. She cannot mediate learning successfully, however, until she understands what her students will be going through as they try to master the key scientific principles and use them to construct explanations. There are two aspects to this understanding: (1) the nature of students' prior knowledge and how it must change, and (2) the strategies that students use for processing new information about science.

Students' Prior Knowledge and the Process of Conceptual Change

Explanations provide a link between science content and students' prior knowledge when facts do not. Students generally just do not know scientific facts before they begin a new unit, so if science is defined as consisting of facts, rules, and definitions, it often appears that students simply know nothing about the topic they are to study. A different and more complex picture emerges, however, if students are asked to explain things that they see in the world around them. It turns out that students *do* have explanations, but not the ones that scientists have developed. A look at the nature of students' explanations reveals why learning science is so complex and difficult for most students.

Consider, for example, the explanation quoted above from Ms. Lane's textbook about how light helps us to see objects. What do students say if they are asked how light helps us to see things? We asked this question to large numbers of fifth graders (Anderson and Smith 1987; Eaton, Anderson, and Smith 1984), and the following are typical answers:

> "In the dark, you can't see anything, so it makes things so you can see them."

"It brightens the path to the object we are looking at."

"It lights up the object so you can see it."

These student explanations certainly are not untrue or unreasonable, but they do not seem entirely satisfactory, either. What is missing (aside from some fancy vocabulary) in these explanations that is present in the textbook explanation quoted above? You might want to take a minute to compare the text and student explanations for yourself before reading on.

* * * *

There are actually several problems with the student explanations. The first one, for instance, is not really an explanation at all; it restates the question in different words. Perhaps the most important thing to notice about these explanations—a characteristic that shows up over and over whenever students discuss vision—is that they *do not mention reflected light*. In our investigations of students' ideas about how we see, we have found that only about 5 percent of the fifth graders we studied were aware that reflected light played a role in vision. The rest thought that light merely made things bright or visible so that we could see them directly.

Our investigations of students' conceptions of light and vision revealed that students' explanations of common phenomena differed from scientific explanations in many other ways as well. Students generally believed that their eyes worked by seeing objects rather than by detecting light, for example. Many students were not sure whether light was always in motion or not. Most students thought that white light was clear or colorless, rather than a combination of the colors of the spectrum. We discovered that students had an interconnected network of mutually supporting beliefs about light and vision that made sense to them, but that were incompatible with scientific ideas presented in the textbook.

The books and review articles cited at the beginning of this chapter describe similar student beliefs about dozens of other scientific topics. These student beliefs are incompatible with scientific theories; for this reason we label them *misconceptions*. They are not, however, foolish or completely wrong. They are often compatible with our normal language. (For example, we generally say, "I look out the window and see the tree," rather than, "I see the light reflected by the tree coming in the window to my eyes.") Misconceptions are generally ideas that are reasonable and appropriate in a limited context, but students inappropriately apply them to situations where they do not work.

A strategic teacher in science must understand the dual nature of students' prior scientific knowledge. In part, prior knowledge is the foundation for meaningful learning in science. Students must relate scientific theories to their own ideas about the world in order to see science as a means of understanding the world rather than a collection of arcane and disconnected facts. On the other hand, students' misconceptions are barriers to successful learning. They must give up or modify many of their beliefs about the world in order to truly understand science.

Thus, science learning is a complex process of *conceptual change*, in which students must modify some of their beliefs about how the world works while strengthening and reorganizing others. In this way, successful learners of science gradually reconstruct their understanding of the natural world. The interconnected and mutually supporting nature of student misconceptions makes this process an arduous one for most students.

Students' Science Learning Strategies

Ms. Lane thinks about the process of conceptual change, and it helps her understand why she was so concerned about her students' learning. They were not really going through the process of conceptual change. They were just memorizing a few facts without altering their misconceptions at all! Still, she is puzzled. She can find many passages in the book like the one about vision quoted earlier. How could the students *read* those passages without seeing the conflict?

My colleague Kathleen Roth has conducted a study that helps to answer Ms. Lane's question (Roth 1985). She conducted careful analyses of students' thinking as they read science textbooks. The students in her study were middle school students reading chapters from one of three texts, two commercial and one experimental, about photosynthesis, the process by which plants use sunlight to make their own food.

In order to truly understand these textbooks, the students in Roth's study had to go through a process of conceptual change. All the students in the study began with a number of important misconceptions, including the belief that plants get their food from the soil rather than making it themselves. She found, however, that most students did not change their basic ideas about food for plants. Instead, they relied on reading strategies that enabled them to cope with the normal demands of school through superficial learning rather than true conceptual change. Roth described at least five different strategies or approaches to reading science textbooks, only one of which resulted in conceptual change. Those five strategies are described below (from Roth 1985).

1. Overreliance on prior knowledge in order to complete a school task. Students using this strategy were generally below-grade-level readers who interpreted the text almost completely in terms of their prior knowledge and incorrect beliefs about plants and food. When asked to recall what the text said, for example, they frequently attributed to the text things that were not in the text but came from their prior knowledge. Although they reported the text made sense to them, these students appeared to avoid thinking about the text itself as much as possible. If they could decode the words and get enough of the gist of the text to call up some appropriate prior knowledge, they reported that the text "made sense."

For example, Maria read a section of the *Concepts in Science* (Brandwein et al. 1980) text that used milk as an example of how all food can ultimately be traced back to green plants, the food producers. Maria announced that "most of this stuff I already knew," and that this was the easiest section to understand. "It was about milk." When probed, she expanded her summary of the text: "It's just about milk...how we get our milk from cows." She never picked up any notion that plants make food. This is typical of her pattern of reading to find similar ideas, ignoring the rest of the text and relying on prior knowledge to fill in the details.

The students using this strategy answered questions posed in the text by thinking about their real-world knowledge about plants rather than using text knowledge. Without thinking about plants' roles in producing food, for example, Maria came up with the right answer to the following question by thinking about her prior knowledge:

> Question: All the food we eat can be traced finally back to the
>
> (a) green plants
>
> (b) cows

Maria correctly picked (a) and explained: "I don't know...I just circled green plants because everybody eats...not everybody eats cows but *everybody* eats green plants."

Thus strategy 1 enabled students to complete assigned tasks and comply with school expectations without really engaging the text at all.

2. Overreliance on words in the text in order to complete a school task. Another group of students, also generally poor readers, isolated words without relating those words to each other or to any prior real-world knowledge. In recalling what they had read, these students identified words or phrases ("It was about chlor-something and an ecosystem") without giving any description or meaning to them. In spite of this lack of attention to meaning, these students reported that they "understood" the text if they were able to decode the words and to identify details in the text that satisfactorily answered questions posed by the text. They reported feeling confused only when they encountered unfamiliar words that they could not decode. For example, when asked whether there were any places in the *Modern Science* text (Smith, Blecha, and Pless 1974) that were confusing, Tracey reported that it was just "some of the words I didn't get." On day three, she pointed out the following *words* as places where she was confused: fermentation, chlorophyll, chloroplast, cotyledon, embryo, dormant.

In answering text-posed questions, students using this strategy simply looked for a "big" word in the question, located that word in the text, and copied the word along with words surrounding it in the text. These copied words may or may not have sensibly answered the question, but the students were satisfied just to have an answer. Frequently, this strategy produced answers that would be acceptable to most teachers. When students were asked interview questions about real-world plants, students relied totally on their experience and prior knowledge. They saw no relationship between the text and the plants. Thus, Tracey recalled the book being about "chlorophyll" and "photosynthesis," but these words were never mentioned when she was asked how a particular plant gets its food.

The students' use of the text was driven by their school knowledge of what was needed to finish the work. They had found that being able to recite key words and phrases from the text (especially large vocabulary words) can often get you by. Real-world knowledge was for them a completely separate realm of knowledge used to explain everyday, nontext phenomena.

3. Overreliance on unrelated facts in the text due to an "addition" notion of learning. This strategy was used by better readers who tried to learn from the text instead of just trying to finish the assigned work. These students held the view that school science learning is all about developing lists of facts about natural phenomena. Their prior experience with schooling had convinced them that memorization of unrelated facts is satisfactory learning. They rarely attempted to relate the facts from the text to each other or to their real-world knowledge about plants.

Students using this strategy often had fairly accurate and complete recall of explicit text material. However, they remembered ideas in no particular conceptual order, they placed equal emphasis on trivial details and on main concepts, and they did not link facts together to develop an overall picture of the main concepts. For example, Myra remembered a lot of details about an experiment that had been described in the *Concepts in Science* text:

> She had some fish and she had some plants in there and one day she was looking at them and a bubble came out of one of the plants. And she started experimenting a little, and she noticed they were giving off oxygen...They asked us what we think about what is she trying—is it oxygen, they asked us what we thought. I put one time it did and one time it didn't...They said the first time it wasn't sunny all the time. The first time it was out for one week and every day it was sunny.

However, when the interviewer asked Myra whether the girl doing the experiment had made a conclusion about the role of the sun, Myra said simply, "no." Although she remembered a lot of details, she missed the critical reason for including the experiment in the text.

Like students using the first two strategies, students using this strategy answered questions about real plants without making reference to any of the facts they had read about in the text and included in their recall. Since they separated school science from plants in the real world, they could not use the information in the text to change their misconceptions about food for plants. Thus the text failed to help them through the process of conceptual change.

4. Overreliance on prior knowledge in order to make sense of disciplinary/text knowledge. This strategy was used by a relatively large number of students, most of whom were reading at or above grade level as measured by standardized achievement tests. These students did not focus primarily on developing strategies to get by in school. Instead, they seemed to be genuinely trying to make sense of the text and the disciplinary knowledge it contained. Thus, in contrast with the first three strategies, this is a sophisticated strategy in which readers attempted to link prior knowledge and text knowledge. These students generally expected the text to *confirm* their prior knowledge, however, so that their goal in reading the text was basically to verify and add details to what they already knew. This attitude was expressed by some as, "Basically, I already knew all this."

Because the students' prior knowledge was so strongly held and because it was often in conflict with the text content, the students using this strategy had to distort or ignore some of the text information to make it fit. Thus, these students *did* integrate prior knowledge and disciplinary knowledge. However, with prior knowledge taking the driver's seat in the process, learning was often quite different from what was intended by the authors of the text.

5. Conceptual change strategy. Some students in Roth's study, all but one of whom were reading the experimental text, used a conceptual change reading strategy. These students also worked to reconcile text information with their prior knowledge, but they allowed the text to take the driver's seat in their attempts to integrate real-world knowledge and text knowledge. Thus, they used text knowledge to *change* their ideas about food for real plants. These students recognized the conflicts between what the text was saying and their own naive theories, and this conflict was resolved by abandoning or changing their misconceptions in favor of the more powerful, sensible disciplinary explanation.

All of the students using this strategy displayed abilities as they read the texts that were not observed among the other students, including the following:

(a) They could recognize and state the main concepts of the text.

(b) They were aware of the *conflict* between text explanations and their misconceptions and willing to abandon misconceptions to resolve the conflict.

(c) They were aware that the text was leading to changes in their own thinking about real-world knowledge.

(d) They were aware of places where the text explanations were confusing because they were in conflict with their previous beliefs.

(e) They could use text ideas to answer questions about real plants.

The students using the conceptual change strategy found it to be hard work; they were the *most* likely to acknowledge feeling confused or having difficulty understanding the text. All of the other strategies left the students feeling confident that they had "understood" because they never really recognized the differences between the text's view of how plants get their food and their own naive views. All of the reading strategies "worked" in the sense that they helped students get through school tasks, but only the conceptual change strategy truly leads to learning with understanding.

Implications of Research on Student Learning

As Ms. Lane looks at her text she begins to see why most of her students do not understand it very well. The text does contain some useful explanations, but it does not really challenge students to think about or change their own explanations. It uses boldface type to emphasize new words, not important ideas. The "main ideas" listed for each chapter are generally facts. The questions in the book and unit test can be answered from students' prior knowledge or through factual recall. No wonder students generally use ineffective strategies when they read science texts!

Ms. Lane, though, does not just want to explain her students' difficulties. She wants to help them understand science. What can she do? It would be nice if she could get a better text. (In Roth's study, the conceptual change reading strategy was used by six of the seven students reading the experimental text, but by only one of the 12 students using the two commercial texts, a seventh grader reading at the twelfth-grade level.)

Like most teachers, however, Ms. Lane is stuck with the text she has. What can she do in her classroom to make cognitive instruction a reality? How can she, as a teacher, help her students engage in conceptual change learning? These are instructional questions, and they are addressed in the next section.

The Instructional Problem: Helping Students Understand

This section summarizes research findings about some of the key elements of a teaching strategy that leads to conceptual change in students. The research on which this section is based includes work by Minstrell (1984) and Nussbaum and Novick (1982) as well as our own research (e.g., Roth, Anderson, and Smith forthcoming). The research is built on comparisons between teaching that does lead to conceptual change in most students with teaching of similar materials under similar circumstances that does not. Thus, the techniques recommended here have been used successfully by classroom teachers (including Ms. Lane). This research is reviewed in Anderson and Smith (1987).

Teaching for conceptual change can be described as a three stage-process. First, there is a *preparation* phase, in which students begin to think about the phenomena that will be explained in the unit, discuss their own explanations, and become aware of the limitations of their naive explanations. Second, there is a *presentation* phase, in which teachers explain key scientific principles and theories. Finally, there is a phase of *application and integration*, in which students apply the scientific principles to new phenomena and integrate those principles and theories into their personal and scientific knowledge. These three phases are summarized in the Planning Guide for Teaching for Conceptual Change at the end of this chapter and are discussed below.

Preparation for Conceptual Change

Students generally begin learning about a topic with, at best, a vague sense that they could learn more than they now know. They generally expect that learning to consist of facts, rules, and definitions, and they generally begin with few unanswered questions about the topic. They seldom have really noticed and observed the phenomena associated with that topic. Many of Ms. Lane's students, for example, have never noticed where they have to stand to see someone else in a mirror or what a pencil looks like in a glass of water. They have never asked themselves where the colors of the rainbow come from or which direction light is traveling when they "look out" of a window. If students are to engage successfully in the active, difficult work of conceptual change learning, they need a sense of what the *potential* is for their learning, of what it is they don't know about they topic but would like to find out.

A variety of instructional techniques can help students become aware of potential learning and develop personally owned questions that will drive active inquiry into a topic. Simple advance organizers that provide a conceptual overview of what students will learn are often useful. It is also important to get students engaged in observing and talking about the phenomena that the unit will focus on. Encouraging observation of discrepant events, where students see things happen that their own conceptions do not lead them to expect, is useful. It is also useful to promote careful observation and thinking about everyday events or objects that play a significant role in students' daily lives, especially if you can ask provocative questions that will lead students to question how they understand those objects and events. Dialogue and debate among students can help them become aware of other ways than their own to think about the objects and events that they are observing. All of these techniques help students become actively engaged with the content of the unit.

Teachers can also learn a great deal from these initial discussions. By listening to students' explanations and ideas, they can begin to gain some insight into their students' prior knowledge and misconceptions. These insights will be of great use to them as they continue teaching the unit.

Introduction of Scientific Concepts

While such exploratory activities are useful, they do not in themselves lead to conceptual change learning (cf. Smith and Anderson 1984). Left to their own devices, students may discover many interesting things about plants or light, but they will develop scientific ideas about photosynthesis or vision about as rapidly as the human race: in other words, not in a single lifetime.

Therefore, scientific concepts need to be explicitly introduced and taught to students. Students must see the scientific concepts as important, meaningful, and comprehensible. They must understand that the scientific concepts are different from their own previous ideas, but not so strange or difficult that they cannot be understood.

We have found that students comprehend scientific concepts better when they are introduced in the context of a meaningful problem, not simply as facts or definitions to be learned. Thus, for example, the previously quoted passage in Ms. Lane's text explaining how we see will be more meaningful to students if they have been discussing the question before reading the passage.

Contrasting scientific concepts and common student misconceptions is also useful. Ms. Lane, for instance, can help her students by clearly pointing out that their explanations of how we see do not mention reflected light, but the book does; or she can help them understand that looking out of a window is actually detecting light coming *in*.

Finally, it is important to emphasize key concepts through such devices as repetition and highlighting (using bold print and putting key concepts in boxes). This may seem like obvious advice, but look at what is emphasized in a typical science textbook: the vocabulary words!

Application and Integration

Research on learning indicates that students initially learn new concepts in specific task contexts and often have difficulty transferring what they have learned to other contexts. This is true even if the concepts are in fact broadly and generally applicable. Thus students cannot fully master a new concept or appreciate its significance until they have successfully used it in a variety of different contexts.

Therefore each important concept should be included in a variety of different tasks, and its presence and importance should be explicitly signaled. For example, the idea that seeing is detecting reflected light can be used to explain why we cannot see through walls (do they block our "view" or the light coming toward us from the other side?), how our eyes work, how we see colors, why frosted glass makes things look blurry, and so forth. Only when students have successfully explained these and other situations can they truly be said to understand that we see by detecting reflected light.

Discussion of Instructional Strategies

One useful way of bringing together the teaching strategies discussed above is to consider a science classroom as a *learning community*. An ideal community for the learning of science needs to be different from a typical science classroom. Students should be actively engaged not just in learning facts and practicing skills, but in practicing the activities of scientifically literate adults: explanation, description, prediction, and control of objects and events in the natural world. Students in such an ideal learning community learn science from sources of *authority* such as textbooks and the teacher, from *evidence* that they acquire by working with natural objects and events, and from *communication* with each other and their teacher.

Creating such an ideal learning community in an ordinary classroom is not an easy task. Textbooks and other teaching materials are generally designed to support memorization and skill practice, not the understanding of science. Young children are typically curious about the world, but they are not naturally disciplined learners of science. Older students often habitually use learning strategies that help them get through science classes, but that does not lead to true scientific understanding. Teachers often work under conditions that do not encourage the creation of ideal learning communities. Nevertheless, we have seen Ms. Lane and other teachers use teaching strategies like those described above to create true (if not ideal) learning communities in their classrooms. They are truly engaged in cognitive instruction in science.

Conclusion

In discussing the problems that Ms. Lane and other teachers encounter in planning and teaching science, I have implicitly answered the questions posed at the beginning of the chapter. This concluding section makes those answers more explicit.

1. Is the framework for strategic teaching and learning developed in the first half of this book consistent with research in science? Yes, it is. I have tried to show in this chapter how successful learning and instruction in science has each of the characteristics of successful learning and instruction discussed in Part I of this book. Moreover, I have tried to portray strategic teaching in science as a complex and demanding enterprise. It is hard work, and it demands an understanding of science content, of students and their scientific thinking, and of appropriate teaching strategies. There are at least three ways in which we can work to help science teachers become more successful at strategic teaching in science.

First, we can improve science textbooks and other teaching materials. Roth and others have demonstrated in their research that it is possible to write textbooks that lead to much higher levels of student understanding than current commercial textbooks. Commercial publishers, however, produce what the market demands, which is currently broad content coverage, lots of facts, and lots of features. The true impetus for improvement in textbooks must come from the buyers, including members of the Association for Supervision and Curriculum Development.

Second, we can improve preservice and inservice teacher education to help teachers become aware of the issues discussed in this chapter, master cognitive instructional teaching strategies, and ask questions about their teaching that will lead to continued professional growth and development.

Finally, we can improve the conditions of teaching. Strategic teaching can become a widespread reality in the schools only if teachers have time for planning, grading, and professional growth, if they are rewarded for trying and using cognitive instructional strategies, and only if communities are formed in which teachers can help each other think about and try to improve their science teaching.

2. To what extent is the framework developed in Part I of this book helpful for planning sequences of instruction in science? I think it is very helpful, especially the generic planning guides: Planning Guide 1 for Thinking Processes and Planning Guide 3 for Strategic Teaching. The Planning Guide for Teaching for Conceptual Change that I developed for this chapter embodies the concepts in both generic guides. Specifically, the Planning Guide for Teaching for Conceptual Change conceptualizes science learning and instruction in terms of the three-phase process just like that used for other subjects, as well as particular teaching/learning strategies emphasized in Part I such as linking new information to prior knowledge.

3. What adaptions, if any, need to be made to teach low- and high-achieving students so that all students will benefit from cognitive instruction? I do not have a fully adequate answer to this question. In our studies we have expanded the percentage of students' understanding from the 0-20 percent range to the 50-80 percent range. This is clearly a great improvement, but we have yet to develop techniques that reach the lowest-achieving students. In particular, we suspect that we are not reaching (a) passive learners, (b) students who are frequently absent, (c) students who are learning-disabled or handicapped, and (d) students who are severely deficient in reading and writing skills.

I feel confident, however, that the path to improvement does not lie in special programs, particularly for low-achieving students. Instead, we need to find ways of creating classroom learning communities in which all students are full participants. Low-achieving students need very much to practice expressing their ideas, to listen to the ideas of others, and to have others listen to them.

This is one of many problems that I am still wondering about and working on as I look for better ways of teaching science. There is a sense in which strategic teaching and

Planning Guide for Teaching for Conceptual Change

Student Thinking	*Teaching Strategies*
Preparation for Conceptual Change	
Anticipate learning to take place, Develop adequate descriptions of natural phenomena, Develop awareness of, and dissatisfaction with, own explanations	Provide advance organizers, Generate observation, discussion and writing about everyday objects and events, Question and debate explanations, Generate observation and discussion of discrepant events
Introduction of Scientific Conceptions	
Achieve initial minimal understanding of scientific explanations, Understand scientific conceptions as reasonable alternatives to own reasoning (not too difficult to understand, not just additions to their own ideas)	Emphasize key principles and theories, Contrast misconceptions and goal conceptions, Introduce conceptions in the context of meaningful tasks
Application and Integration	
Understand scientific principles and theories as widely applicable, Understand interconnections with other personal and scientific ideas	Explicitly signaled inclusion of conceptions in other tasks, especially --tasks in everyday contexts, --tasks in other scientific contexts

cognitive instruction in science are ideals that neither I nor any other science teacher can ever fully achieve. This chapter, however, describes many specific actions that teachers, administrators, and others can take to improve science teaching. Long-term improvements in science education will arise from the accumulation of those specific actions.

References

Anderson, C.W. "Improving College Science Teaching: Problems of Conceptual Change and Instructors' Knowledge." Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, 1986.

Anderson, C.W., T.H. Sheldon, and J. DuBay. "The Effects of Instruction on College Non-majors' Conceptions of Respiration and Photosynthesis." *Journal of Research in Science Thinking* (forthcoming).

Anderson, C.W., and E.L. Smith. "Teaching Science." In *The Educators' Handbook: A Research Perspective*, edited by V. Koehler. New York: Longman, 1987.

Blecha, M.K., P.C. Gega, and M. Green. *Green Book. 2nd ed.* Exploring Science Series. River Forest, Ill.: Laidlaw, 1979.

Brandwein, P., E. Cooper, P. Blackwood, N. Cotton-Winslow, M. Giddings, E. Romero, and A. Carin. *Concepts in Science.* New York: Harcourt, Brace, Jovanovich, 1980.

Carey, S. "Cognitive Science and Science Education." *American Psychologist 41*, 10 (1986): 1123-1130.

Clement, J. "Students' Preconceptions in Introductory Physics." *American Journal of Physics 50*, 1 (1982): 66-71.

Driver, R., E. Guesne, and A. Tiberghien, eds. *Children's Ideas in Science.* Philadelphia: Open University Press, 1985.

Eaton, J.F., C.W. Anderson, and E.L. Smith. "Students' Misconceptions Interfere with Science Learning: Case Studies of Fifth-grade students." *Elementary School Journal 84*, 4 (1984): 365-379.

Hollon, R.E., and C.W. Anderson. "Teachers' Understanding of Students' Scientific Thinking: Its Influence on Planning and Teaching." Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, 1986.

Kuhn, T.S., *The Structure of Scientific Revolutions.* 2nd ed. Chicago: University of Chicago Press, 1970.

Mayr, E. *The Growth of Biological Thought.* Cambridge, Mass.: Belknap, 1982.

Minstrell, James. "Teaching for the Development of Understanding of Ideas. Forces on Moving Objects." In *Observing Science Classrooms: Perspectives from Research and Practice, 1984 Yearbook of the Association for the Education of Teachers in Science*, edited by C.W. Anderson. Columbus, OH: ERIC/SMEAC, 1984.

Nussbaum, J. and S. Novick. "Alternative Frameworks, Conceptual Conflict, and Accommodation: Toward a Principled Teaching Strategy." *Instructional Science 11*, 3 (1982): 183-200.

Roth, K.J. "Using Classroom Observations to Improve Science Teaching And Curriculum Materials." In *Observing Science Classrooms: Perspectives from Research and Practice, 1984 Yearbook of the Association for the Education of Teachers in Science*, edited by C.W. Anderson. Columbus, OH: ERIC/SMEAC, 1984.

Roth, K.J. "Conceptual Change Learning and Student Processing of Science Texts." Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1985.

Roth, K.J., C.W. Anderson, and E.L. Smith. "Curriculum Materials, Teacher Talk, and Student Learning: Case Studies of Fifth-Grade Science Teaching." *Journal of Curriculum Studies* (forthcoming).

Smith, E.L., and C.W. Anderson. "Plants as Producers: A Case Study of Elementary Science Teaching." *Journal of Research in Science Teaching 21* (1984): 685-698.

Smith, H.A., M.K. Blecha, and H. Pless. *Modern Science, Level Six.* River Forest, Ill.: Laidlaw, 1974.

Toulmin, S. *Foresight and Understanding.* Great Britain: Anchor Press, 1961.

West, L.H.T. and A.L. Pines. *Cognitive Structure and Conceptual Change.* Orlando: Academic Press, 1985.

TEACHING SCIENCE FOR UNDERSTANDING

James A. Minstrell

Early in my experience as a high school science teacher, I became concerned about my ineffectiveness at teaching students to transfer their understanding. I don't mean that I was a bad teacher. My students tested well, and my administrators gave me glowing evaluations, but I was amazed at the relatively little effect I had on my students' understanding of the ideas of physics.

In the early 1970s, I, like others, thought the difficulty might be my students' lack of ability with formal operational reasoning. However, in my investigations of students' operational reasoning capacity, I found that the results of reasoning tests depended on students' conceptual knowledge. For example, consider a task involving two equal-sized balls of clay. After establishing that the two balls weigh the same, one of them is flattened into the shape of a pancake. Students are again asked to compare the weights. The task tests students' understanding of the operation of conservation, in this case conservation of weight. Many students "fail" the task not because they can't do conservation reasoning, but because their concept of weight or gravity involves air pressure, and a flattened ball of clay has more upper surface on which "air presses down." This represents a conceptual error rather than an operational reasoning error. Students are bringing to the situation content ideas that greatly affect their performance on questions that are supposed to test their operational reasoning. Even on problems that require the same abstract reasoning on similar quantitative data, the specific content of the situation affects the reasoning strategies students use.

The growing research on students' conceptions in the late 1970s helped me become more effective at teaching for the understanding of ideas, especially on problems requiring transfer of conceptual understanding to novel situations. In this chapter I describe how my experience in the classroom and the research on cognitive processes and related areas now guide my teaching. My teaching is more informed, and I have specific conceptual goals, based on recent information about the cognitive nature of the learner and the nature of the learning process.

---

*This article first appeared in Lauren Resnick and Leopold E. Klopfer, Eds. (1989), **Toward the Thinking Curriculum: Current Cognitive Research**, 1989 ASCD Yearbook, pp. 129-149. Alexandria, VA: Association for Supervision and Curriculum Development. Copyright © 1989 by ASCD. Reprinted by permission.*

The first section includes annotated descriptions of a lesson about force, a central concept in Newtonian physics and one of the topics in physics learning that has been heavily studied by cognitively oriented researchers. The next section gives further examples from two additional lessons that illustrate the care needed in designing and implementing instruction aimed at fostering conceptual restructuring in the science classroom. Finally, I discuss and summarize the major instructional principles to be drawn from cognitive research on science learning. Although all the examples in this chapter are from physics, my colleagues assure me that the principles apply across the sciences and beyond.

An Introductory Lesson on Force

Students come to the classroom with initial conceptions organized by their experiences. While it is probably true that each individual has a unique set of past experiences, there is a remarkable consistency among the naive conceptions about motion that students in various cultures bring to the physics classroom.

Students' initial ideas about mechanics are like strands of yarn, some unconnected, some loosely interwoven. The act of instruction can be viewed as helping the student unravel individual strands of belief, label them, and then weave them into a fabric of more complete understanding. An important point is that later understanding can be constructed, to a considerable extent, from earlier beliefs. Sometimes new strands of belief are introduced, but rarely is an earlier belief pulled out and replaced. Rather than denying the relevancy of a belief, teachers might do better by helping students differentiate their present ideas from and integrate them into conceptual beliefs more like those of scientists.

Another goal for teachers is understanding how new experiences and ideas can be designed to help students make sense of their physical world. Cognitive research emphasizes that instruction is not simply telling students to stop thinking the way they have been and think according to the example of the teacher, text, or some other authority. Note that the discussions in this lesson are rooted in common, concrete experiences. When phenomena come from everyday experience, they are more difficult to deny. Students respond better when information is linked to something they've experienced rather than to a pronouncement from authority.

Several additional kinds of research influence this lesson. Research on verbal interaction in the classroom helps me to define the climate that fosters concept development. Related areas of investigation, such as brain research and learning styles, are mentioned in passing largely because, in my experience, they appear relevant. Their actual relation to concept development is yet to be understood in an operational sense.

This first lesson is presented below as a transcript interspersed with text that explains the research-based rationale behind the instructor's decisions. The teacher's primary goal is to give students an idea of the content in the unit on the concept of force. While being open to and accepting of what students say, the teacher attempts to help students clarify their initial ideas about force. Also, the teacher assesses students' understanding of related ideas and builds the semantic network of ideas the unit activities should uncover. Thus, the curriculum is not predetermined but can be built, depending on conceptual and rational needs and, to some extent, the contexts of interest of the students.

Teacher: *Today we are going to try to explain some rather ordinary events that you might see almost any day. You will find that you already have many good ideas that will help explain those events. We will find that some of our ideas are similar to those of the scientist, but in other cases our ideas might be different. When we are finished with this unit, I expect that we will have a much clearer idea of how scientists explain those events, and I know that you will feel more comfortable about your explanations.*

Students' views of the physical world, although valuable in their daily world, differ from the scientists' views of the world (Driver 1983, Gilbert and Watts 1983).

Teacher: *A key idea that we are going to use is the idea of force. What does the idea of force mean to you?*

(A discussion follows. My experience suggests that the teacher should allow this initial sharing of ideas to be very open.)

To students, the word *force* has multiple meanings (Viennot 1979). Typically, they suggest it means something to make something happen, to create change (Minstrell and Stimpson 1986). Their descriptions often include related words: energy, momentum, strength, fast moving, or a hit (Osborne 1984). The terms "push" and "pull" usually are mentioned. Often, discussion stops at this point and the class moves on with both students and teacher believing they have a common understanding of the term. But contrary to the scientist's conception, students believe that the push or pull must be due to some causal agent. This agent, in their view, makes things happen; it is the motivator or the energy source in the situation (Minstrell and Stimpson 1986).

Teacher: *You've mentioned words that represent many ideas. Most of them are closely related to the scientist's idea of force, but they also have meanings different from the scientist's ideas. Of the ones mentioned, probably the one that comes closest to the meaning the physicist has is the idea of push or pull, so we'll start with that. We'll probably find out that even that has a slightly different meaning to the physicist. (The teacher should allow the class to begin with this meaning for force rather than present an elaborate operational definition.)*

Through the course of events in the classroom, the students' ideas will evolve. Lessons have specific purposes, so students gradually differentiate and then recombine their initial ideas into a view that is more logically consistent when applied across a broad range of situations. For example, certain lessons will help students integrate passive support by a table and active support by a hand under the same general concept of force (Clement 1987; Minstrell 1982b). Students will learn to differentiate actions like force, impulse (force x time), and work (force x distance) from each other and from properties (or states) of objects or systems—properties like velocity, acceleration, momentum, and energy (McDermott 1984, Minstrell and Stimpson 1986). The resulting conceptions will probably be more like the scientist's, and they will have grown out of pieces of knowledge that make sense to students (diSessa 1987).

Teacher (dropping a rock): *Here is a fairly ordinary event. We see something like this happening every day. How would you explain this event, using your present ideas about force?*

The teacher picks an ordinary occurrence, which is preferable to a special situation for instruction. If we choose ideal case situations, especially in physics, we run the risk that students will think of physics as episodic, dependent on specific situations, and separate from daily experiences with the physical world. Instead, students should consider physics in the context of their everyday lives. Part of the value of the formal ideas of physics is that they apply consistently across common phenomena.

Teacher: *Don't speak right now. I want each person to have a chance to do his or her own thinking.*

To the extent possible, the teacher actively engages each student in the activity. Students have a vested interest in the outcome of the activity if they have first committed themselves to it (Minstrell 1982b, Rowe 1974).

Teacher: *Make a drawing of the situation and show the major forces acting on the rock when it falls. Use arrows to represent the forces, and label each as to what exerts the force.*

The teacher encourages students to represent the situation. Popularizers of brain research, like Buzan (1974), believe that representing situations with pictures as well as words activates the visual and verbal parts of the brain. It is suggested that this enhances access to knowledge by creating linkages to more parts of the brain. This makes sense to me as an experienced teacher and learner. If people have recorded their ideas on paper, they are less likely to deny the existence of their initial ideas and they are more likely to notice inconsistencies between their original idea and what actually happens (Loftus 1980).

Students (naming the forces they have represented):

Gravity by the Earth.

Weight of the rock.

Both gravity and the weight.

Many students think of weight as a property of a body and have not conceptualized the weight of a body as the force on that body exerted by the Earth, basically a relationship between the object and the Earth (Carey 1987).

Students: *Air*

Some students mean air pressure in the downward direction. Air pressure is apparently not distinguished from gravity (Minstrell 1982a). For others, air is the main effect that makes the weight of the object, and the gravity effect is just 1/6 the weight, i.e., the residual left after air is removed. This was apparently deduced from the fact that on the moon things weigh 1/6 of their weight on Earth (Jung 1984). (The relation between gravity and air pressure is the subject of another lesson and not presented in this chapter.)

Students: *Resistance.*

Friction.

Resistive force, while common to everyday experience, involves rather complex mechanisms. To many students, friction does not act in a particular direction. It is somehow just present in situations (Clement 1987). Due to the complexities of both the scientist's and the students' views, the instructor decides not to center the discussion around this idea until some more basic notions about force are clarified.

Students: *The spin of the Earth.*

Nuclear forces.

Since many of these ideas are complicated and rely on a more basic understanding of force, the teacher attempts to limit the discussion.

Teacher: *Which of these is the major force, or which are the major forces acting on the rock while it is falling?*

Students: *Its weight.*

Gravity.

If the weight/gravity relation hasn't been developed in an earlier unit, this would be a good opportunity to do so (Champagne, Klopfer, Solomon, and Cahn 1980).

Teacher: *Actually, weight is the name we give to the gravitational pull on an object near the Earth. So, weight and gravity really represent the same idea. Let's take this as an assumption for now. We really should come back later and develop that more fully.*

Sometimes the teacher needs to make the choice to state a piece of information and ask students to take that as a basic assumption, so the argument can continue. That is what the teacher does in this instance. But if the distinction was developed earlier in class, the students could now be encouraged to briefly review those arguments. Teaching experience suggests that the use of earlier arguments helps students make them part of their general long-term knowledge (Minstrell 1984). Learning research supports that experience (Buzan 1974).

Teacher: *Is the falling rock moving at a constant speed, or is it speeding up or slowing down? How do you know?*

The students have had prior class experience discriminating between constant uniform speed on the one hand and changing speed on the other. This lesson provides an opportunity for the teacher and students to revisit the arguments for separating those motions. While one of the main objectives of a later lesson will be to differentiate the force acting on the object from the resulting action of the object, the teacher introduces that idea here by asking about the motion as a separate question from the forces. Also, it is important to have students attempt to justify their inferences as well as make them. "How do you know?" "How did you decide?" "Why do you believe that?" These are all useful examples of justification of knowledge questions (Arons 1983, Minstrell 1987). The teacher should allow several seconds to pass before calling on students or allowing them to answer so everyone has an opportunity to develop an answer. Rowe (1974) calls this period "wait time."

Student: *The same speed all the way because I saw a film where some guy said that all things fall at the same speed.*

The teacher pauses here for three to five seconds before commenting or calling on another student so the first student can evaluate her answer (Rowe 1974).

Student: *No wait, that's if two things fall, they both fall equally fast. I don't know. (More wait time.)*

Student: *I think the rock speeds up.*

Teacher: *What evidence do you have that makes you think it accelerates?*

Student: *The higher you drop it from, the harder it hits the ground or whatever.*

The teacher allows more wait time here, too. Even if students' answers are consistent with the scientist's view, give them time to evaluate their own ideas to decide their validity. The sources of validity should be experience, both informal and experimental, and logical, rational argument, but not the authority of a teacher or a text.

Student: *Also, when one of the groups was studying velocity and acceleration, they dropped a thing and it accelerated.*

Teacher: *So what would a diagram of forces look like to explain the falling rock?*

Student: *Just gravity or whatever down and a little bit of air all around the rock. Like this, I think.*

Student: *There are some of those other forces, but probably the gravity or weight pulling down is the main one.*

Teacher: *Suppose I make the rock speed up slowly in the upward direction, while keeping it in my hand. What are the forces that explain this situation?*

Student: *The hand making the up force.*

Teacher (after about five seconds wait): *Are there any other major forces in this situation?*

Student: *Well, yeah, the weight is still there. That's in the down direction, but the hand force has to be bigger.*

Teacher: *Do others of you agree with her analysis? Anybody disagree?*

Students: *Seems okay.*

Teacher: *How do you know the hand force must be bigger?*

Student: *Because it goes in that direction, up.*

Student: *If you want something to go in some direction, you have to put a force on it in that direction?*

Student: *You have to put a bigger force in that direction than in the other direction.*

Teacher: *While this rock is speeding up in either direction, the force in that direction needs to be bigger than all the forces in the other direction put together. Is that your idea so far?*

The teacher guides the students to clarify more specifically what they are suggesting. Since the pre-course conception of "force of motion" is probably stronger than their conception that "passive objects exert no force, " the teacher may elect to take up forces on static (stationary) objects first.

Teacher: *During the next several days we will look more closely at the idea we call force. Many of the ideas you've suggested today will be useful, but we may also find that we will want to change some of our notions about force to make them more consistent with the phenomena.*

Additional Lessons: Integration and Differentiation of Ideas

The following examples concern several conceptual difficulties that the teacher can help science students overcome by carefully designing classroom instruction and interactions. The first lesson is designed to help students integrate their existing knowledge with new concepts and make its organization in memory more efficient.

Integrating Concepts of Force

Sometimes students are more parsimonious with conceptual knowledge than they need be. Rather than applying a certain concept or interpretation generally in various contexts, they apply it selectively. Research on students' conceptions of force suggests that they make a big distinction between objects they consider to exhibit active actions such as push or pull and those they see as agents of passive actions such as support or blocking. The active actions they definitely call force, but the passive actions do not count as force (Driver 1983, Minstrell 1982b). The main objective of this lesson is to make a reasonable case for classifying both active and passive actions under the same conceptual heading of force.

Teacher: *Suppose we put this rock on the table. The rock is not moving; it is at rest. What are the forces acting on the rock in that condition? How do you decide? I want each of you to draw and write your own best answers. What makes sense to you? We'll share our answers in just a few minutes.*

When students' ideas are recorded on paper, it is easier for them to identify inconsistencies—and not so easy for them to ignore them.

Teacher: *Okay. Now, I'd like people to share their ideas. What are the forces involved?*

Students usually suggest gravity or weight is a downward force acting on the rock. While some students suggest there is an upward force by the table, there is considerable argument whether a passive object like a table can exert a force. Virtually all the students believe the table is supporting the book, but their notions about force involve something much more active. Since force is a conception invented by humans, the students' ideas cannot be demonstrated to be wrong by performing an experiment and collecting data. The teacher must attempt to guide the students to scientists' rationale for the concept of force, which includes passive actions.

After establishing that the intellectual argument centers on whether it seems reasonable to think of a table as exerting an upward force, the teacher moves to another situation, the rock at rest in the outstretched hand. Nearly every student believes the hand needed to exert a force (Minstrell 1982b). Clement (1987) suggests this is one of those universally believed ideas that can be used as an anchor to build an analogical bridge to the "target situation" (the rock on the table). The hand clearly supports the rock, and students also agree that the hand exerts an upward force.

The teacher then invites the class to consider other examples between the commonly agreed upon anchor situation (rock on the hand) and the target situation (rock on the table). For example, the rock might be placed on, or hung from, a spring. It appears that observing the adjustments of the spring to the downward force exerted by the rock is enough to convince the students that the spring also exerts an upward force on the rock, similar to how the muscles in the arm support a rock on a hand.

Some of the students are beginning to ask their own questions about their understanding and that of their fellow students. Research suggests that before students go through conceptual change, they need to see that the new conception is plausible (Posner, Strike, Hewson, and Gertzog 1982). Some seem swayed by logical necessity or convinced by juxtaposing several similar situations (i.e., the rock is at rest and is supported by something in all instances). To these students, that is sufficient for them to invoke the following argument: For the same effect (at-rest condition), we would invoke the same explanation (downward force balanced by upward force). Other students need assistance to see that all the situations are similar (i.e., objects at rest). Some are more influenced by arguments dealing with the physical mechanisms that illustrate how the rock on the rigid, "immovable" table is similar to the rock on the bendable spring. The table, like a stiff spring, will definitely flex, and the amount of bending depends on the weight of the rock.

Teacher: *So, it would appear that even nonliving things that seem rigid do bend ever so slightly. We've observed this ourselves. Because scientists want to be logical and explain the same effect, at rest, with the same reasoning, then all three—the hand, spring, and even the table—are said to exert an upward force. The hand is very active. The table just exerts the force passively, because it happens to be in the way. By these sorts of arguments, scientists have come to believe forces are exerted by active and passive objects. In fact, they believe anything that touches an object must exert some force on the object.*

One of the teacher's main goals for this lesson was to integrate the concept of passive support under the broader concept of force. Students had a notion of passive support, but it was separate from their concept of force. The lesson helped students reorganize initial ideas into a more integrated and logically consistent structure. As a result, the students acquired a more encompassing meaning for the term "force."

Differentiating Related Ideas

Sometimes students have nearly the opposite conceptual difficulty: They use one technical term to encompass too broad a range of meanings. Rather than ignore this problem or simply tell students formal definitions of terms, the teacher can design instruction so students differentiate meanings and organize them into a more powerful arrangement. Consider a lesson directed toward helping beginning physics students differentiate the ideas of force and motion.

The teacher might start with a pre-instructional quiz to identify students' ideas about two situations: a low-friction cart moving at a constant speed across a tabletop and the same cart accelerating uniformly across the tabletop. On quizzes given before instruction, most students explain the acceleration of the cart motion by an increasing force, and they explain the cart moving at a constant speed by a constant "force of motion" in the forward direction. These notions are consistent with daily experience; if we want something to move with a constant velocity, we need to provide a constant push. For naive students, this motion implies force. Whenever they observe or read a description of a moving object, they link a force with the object. They don't really differentiate between the motion of an object and the forces that produce it. They do not differentiate the condition of the object from the actions that caused the condition.

To address the students' conceptual difficulty, the teacher may have them do an experiment where they must consider the two ideas. Students observe both the motion of a cart and the readings on a spring scale that monitors the pulling force. While noting that the cart accelerates across the table, students are surprised to observe that the force-scale reading stays constant. Their conception of a "force of motion" is so strong they will deny the validity of their experimental results. But faced with similar evidence from other groups, they begin to doubt their initial conception and search for resolution. They begin to separate the force acting on the body from the condition of motion of the body.

Then comes the more difficult question: How do we explain the forces on an object moving with a constant velocity? Students have great difficulty coming to the logical conclusion that, to explain the constant velocity of the object in motion horizontally, the horizontal forces must balance each other (Minstrell 1984). To effectively change students' intuition and to have the learning transfer to multiple contexts, they need several more experiences with applying the balanced forces idea. It also seems to take periodic opportunities to repeat the rational arguments used to derive that conclusion.

Teacher: *If we explain the constant acceleration situation with constant net force, what do we do to explain the constant velocity situation?*

Student: *Maybe that just requires less extra force.*

Teacher: *But what happened when we had more or less force in the experiment?*

Student: *We got more or less acceleration.*

Student: *Maybe a constant speed situation could be explained by a constant decreasing force.*

Student: *But we saw that when we had more unbalanced force we got more acceleration, and when we had less unbalanced force we got less acceleration, but we always got acceleration.*

Student: *That's right, but it can't be an increasing force either. That wouldn't make sense.*

Student: *If it can't be constant unbalanced force, and it can't be decreasing force, and it can't be increasing force, what's left? No force?*

The students' laughter illustrates the absurdity of this conclusion (correct from the physicist's view). But to some, the empirical result from the experiment and the logical conclusion from the discussion are beginning to make sense. Yet they still question the value of this idea in the everyday world.

Student: *What about the car? You put your foot down a certain amount on the gas pedal and the car speeds up to a certain speed and then tops out for that pedal setting. If you take your foot off, the car slows down. If you push down farther, it goes faster. That seems like the faster you want to go, the more force you need.*

Students can now use the teacher's guidance and encouragement to analyze the situation involving an accelerating car. The teacher might get them to draw diagrams of the forces on the car at rest, speeding up, at a constant velocity, and slowing down. This is an opportunity for students to use skills and ideas from earlier lessons to identify forces acting on objects in various circumstances. By recalling what they know from earlier class experiences and reviewing the kinds of argument used to arrive at conclusions, they can springboard to a different situation. True, it often is a struggle for beginning students to differentiate between the ideas of force acting on the car and the resulting motion of the car. However, with careful structuring of observational experiences and discussions, the teacher can help students differentiate between related ideas.

Throughout science instruction, there are numerous examples of lack of differentiation by students. For instance, they do not differentiate area from perimeter or volume, or weight from density, size, mass, and volume. To build clearer understanding of science ideas, we need to deliberately design instruction to confront several related ideas, rather than simply tell what each idea means separately.

Summary of Instructional Principles

The two previous sections illustrate how research on cognitive processes and related areas can guide instruction. In this section I want to pull together some of the more important principles from cognitive research that affect science instruction, specifically restructuring knowledge, instructional design, and classroom environment.

Restructuring Knowledge

For many science teachers familiar with the findings of cognitive research, restructuring students' existing knowledge has become the principal goal of instruction. This commitment stems from the realization that the minds of students entering science classrooms are far from blank slates concerning phenomena in the physical world. In physics and in other subjects dealing with day-to-day existence, students come to the classroom with strong initial conceptions (McDermott 1984, Driver, Guesne, and Tiberghien 1985). If instruction does not deal with many of these conceptions, students leave class without having appreciably changed their ideas (Goldberg and McDermott 1987, Minstrell 1984).

As illustrated in the preceding lessons, students' initial concepts are often very different from what we want them to learn. For example, their conception of force involves pushes and pulls by active agents but not by passive ones. Since motion may affect what happens in a situation, students consider it to be a force. To the scientist, motion is a condition or a state of an object, better described by concepts of velocity, momentum, and kinetic energy than by force, which the scientist considers an action on the body. This example suggests that many student ideas are not wrong, but they lack conceptual differentiation and integration. Student knowledge may be constructed from observational data that are similar to scientists' knowledge, but the structures organizing the knowledge are different.

An interesting research question asks: "At what level does one look for a conceptual thought unit?" McCloskey (1984) has represented student thinking about motion at the theory level and suggests it is similar to the impetus theory, an early theory of mechanics. Halloun and Hestenes (1985a, 1985b) center their search at the level of formal physics principles, like Newton's third law. From these perspectives, many student conceptions can be considered wrong. I focus on key words of situational attributes that the student emphasizes, from which student answers can be predicted (Minstrell and Stimpson 1986). This suggests that students may focus on problem features like activity/passivity, causal agency, and motion. DiSessa (1987) searches behind students' words for what he calls "phenomenological primitives." His focus is on constructs based on readily observable phenomena, such as observations of springiness in objects that lead students to classify events in the world as "spring-like."

Consider one more example of basically sound elemental ideas leading to a generally faulty conclusion. From everyday experiences, students come to believe that bigger objects have a bigger downward force acting on them than do lighter objects. Students also argue from the experience of pushing boxes around that if you put a bigger force on an object, it will go faster. Therefore, they conclude that since the heavier object has a greater push downward on it, it falls to the ground in less time. Each piece of knowledge is sound; the logic is sound. Yet this structuring yields the incorrect conclusion that heavier objects fall faster. In this case, there is one more relevant experience that needs to be incorporated: When pushing two boxes that you want to speed up equally, you will need to push the heavier one with a proportionately larger force. Thus, the student's argument, missing one piece, leads him to an incorrect prediction.

As science teachers, we need to decide whether we are going to define students' pre-instructional thinking according to historically developed theories and principles or according to student-described, salient features of phenomena. That decision dictates whether the instructional goal is to produce conceptual theory revolution or to restructure students' use of present knowledge and vocabulary for representing existing knowledge structures. In my teaching, I take the latter approach.

In some cases, students are simply unaware of particular phenomena, in which case the goal of instruction may be to extend their knowledge base. Without any previous exposure to particular phenomena, students seem to organize their observations according to conceptions from other contexts that they perceive to be relevant or analogous. In this chapter, I have focused on two kinds of conceptual transformation that learners go through in the process of restructuring their knowledge. One is knowledge differentiation, which involves splitting apart pieces of knowledge from what initially appeared to be one idea. The second, conceptual coalescence, is the integration of what initially appeared to be separate knowledge pieces into one idea. Just as these processes have been important in the development of ideas through the history of science, they appear equally important in the development of students' knowledge (Carey 1987).

Instructional Design

Cognitive research suggests that certain lesson designs can foster the growth of students' knowledge and understanding. Osborne and Freyberg (1985) have summarized the research on features of instruction that foster conceptual development. In the lessons presented earlier, there was a preliminary phase where the teacher and students identified students' existing ideas. This information might also be discovered through a diagnostic test at the beginning of the course or in a short pre-instruction quiz at the beginning of a unit or just before a major conceptual restructuring activity. Second, there was a focusing phase for students and teacher to clarify the students' initial ideas. This focusing usually takes place during a discussion after students record their initial ideas. The third phase involves an activity or situation that challenges the students' initial ideas. In the "at-rest" lesson, the book on the hand, the book on the spring, and the springy nature of the rigid table constituted the challenge. In the moving cart situation, the challenge came from the laboratory activity that presented results contrary to the students' predictions. Finally, there is an application phase, when students have the opportunity to practice using the new idea in multiple contexts. For our lessons, students needed the opportunity to explain the "at-rest" condition by examining significant horizontal forces when objects are at rest in air, on water, or on an incline. Students needed to be able to explain common moving situations, such as the car accelerating and then moving with a constant velocity, all with the same throttle setting.

Students also need to recycle through the arguments for and against the need for a forward force in later contexts, such as projectile motion and circular motion. The new ideas or restructuring of ideas will need expanded contexts of application. Some students will demonstrate transfer to the new situations; others will revert to previous conceptions. The percent of students successfully transferring to new situations increases with each new episode. The new conceptions are never adopted by all students after only one context, and it is probably unreasonable to assume that the whole class will make the transfer.

In addition to general guidelines for instructional design, research has given us a better understanding of the limitations of perception. Attention is very important. At the beginning of a lesson, we need a question or activity that is thought-provoking or will arouse the learner's curiosity. The use of pre-instruction questions, especially those that ask students to make a prediction that can be readily tested, stimulates interest in subsequent related activities (Jung 1984). Although Bates' (1978) review of the literature, dealing with the value of the laboratory in promoting understanding, found no positive influence, I believe that laboratory activity does have a significant effect if it directly relates to students' initial ideas and allows them to test those ideas (Minstrell 1984). But if the activity is viewed as just one more in a series of unrelated episodes at school, students are not likely to make the connection at all.

It is important to encourage students to seek, identify, and resolve inconsistencies between their ideas and what actually happens in the laboratory or demonstration activity. In the second example lesson, students were asked to compare their explanation for the rock at rest on the hand and on the spring with their explanation for the rock on the table. Aren't all these situations the same? If so, shouldn't all the explanations be similar? If so, how? In considering the forces on the moving cart, some students needed to be guided to notice the difference between their initial ideas and what happened in the laboratory activity. Some lacked respect for their empirical methods and results, and preferred to believe their initial intuitions. They had to be encouraged to check their results with others and then to resolve inconsistencies.

In the classroom, we should strive for transfer as a measure of understanding. Pea (1988) suggests some of the sorts of transfer that can help define understanding. We want to ensure that what students learn in the formal learning context is transferred to everyday life and work situations. In the class discussion during the lesson on forces on moving objects, the teacher encouraged students to consider a common experience: a car speeding up. Students also should have opportunities to acquire and apply knowledge in an integrated manner that approximates the demands of problem solving in everyday life. In the lesson discussion, the newly hypothesized idea about force and previously acquired knowledge were brought together in the common, but complex, situation of the car on the roadway. The teacher did not leave students with only an understanding of the laboratory situation with carts on tables. Conversely, the concepts, skills, and strategies that students have acquired and used effectively outside the school setting should be brought into the classroom and applied to formal learning to demonstrate the value (and limitations) of experience. In the car situation, many of the ideas students brought with them were useful (e.g., wheels spinning on a frictionless surface will not make the car go, a hand out the window feels progressively greater resistance with an increase in speed, and the car's speed does top out if one holds the gas pedal at a certain position).

Finally, students should be able to go beyond the problem originally learned in the classroom and solve related problems, moving in a series of steps to problems that are increasingly different from the learned problem. The key ideas and thinking used in the sample lesson should be evoked again in subsequent contexts, such as a car coasting to a stop or a box sliding to a stop.

To teach for transfer of ideas, we need to include activities requiring students to use new ideas in multiple contexts. To do this we need to offer subsequent contexts, at appropriately spaced time intervals, where students can use the same arguments used in developing the new idea. For example, in the context of the forces on moving objects, students can be asked to explain, perhaps two weeks later, the motion of horizontally launched projectiles. What are the forces acting on the rock after it has left the hand? If the idea of force has been separated from the condition of motion of the object, then students should be able to identify gravity, and perhaps air resistance, as the only significant forces in this context.

Classroom Environment

Teachers need to create an appropriate climate for fostering development of understanding. That climate should encourage questioning. Questions asking for predictions, for clarification of meaning, for justification of how a student decided a particular answer, and for interpretation, explanation, and observations, should prevail. Between questions and answers, as well as after answers, the teacher should allow three- to five-second intervals of wait time to encourage more thoughtful responses (Rowe 1974). There should be an atmosphere of mutual respect among students and between students and teacher. Students bring their ideas into class; since these are the ideas out of which their new ideas will be constructed, the sharing of those ideas should be respected.

Finally, time is a critical factor in developing understanding. There should be enough time within lessons for the inconsistencies illustrated by a new experimental result to register. We must provide the time students need for mental restructuring. Hurrying on to the next lesson or the next topic does not allow for sufficient reflection on the implications of the present lesson. We have found that physics teachers typically spend one to three days developing Newton's laws and the remainder of a two- to three-week unit doing exercises to practice them. When about five days were spent carefully developing the ideas, and even though students didn't get as much practice in manipulating equations, their understanding of the critical ideas at the end of the first or second semester was significantly better, both statistically and educationally. With more careful and extensive development time, more than twice as many students are able to answer difficult conceptual questions (Minstrell 1984).

With an understanding of the research, teachers and developers of curriculum or instruction can be more analytical in their attempts to meet students' needs, better understand why lessons succeed or fail, and gain insight as to what to do or not to do next time.

As a teacher, I would encourage other teachers to become involved in research related to the cognitive processes of their learners. I predict it will enliven their time in the classroom, as it has mine. I also predict it will enhance students', and perhaps their own, understanding of the natural world.

Author's Note: I gratefully acknowledge the financial support of the National Science Foundation in the form of a Presidential Award for Excellence in Science Teaching, part of which was used for the preparation of this chapter. The ideas and words herein are mine alone and do not represent the beliefs of the Foundation. Also, I would like to thank A. Arons and P. Diaz for critically reading earlier drafts of this chapter. Finally, I appreciate the efforts of my students and colleagues in Mercer Island School District toward the improvement of instruction in science.

References

Arons, A.B. (Spring 1983). "Achieving Wider Scientific Literacy." **Daedulus**, 2: 91-122.

Bates, G. (1978). "The Role of the Laboratory in Secondary School Science Programs." In **What Research Says to the Science Teacher**, edited by M. Rowe. Washington, DC: NSTA Publications.

Buzan, T. (1974). **Use Both Sides of Your Brain**. New York: E.P. Dutton.

Carey, S. (1987). **Conceptual Change in Childhood**. Cambridge, MA: MIT Press.

Champagne, A., L. Klopfer, C. Solomon, and A. Cahn. (1980). "**Interactions of Students' Knowledge with Their Comprehension and Design of Science Experiments.**" A monograph published by the Learning Research and Development Center. Pittsburgh, PA: University of Pittsburgh.

Clement, J. (1987). "Overcoming Students' Misconceptions in Physics: The Role of Anchoring Intuitions and Analogical Validity." In **Proceedings of Second International Seminar: Misconceptions and Educational Strategies in Science and Mathematics III**, edited by J. Novak. Ithaca, NY: Cornell University.

diSessa, A. (September 1987). "Toward an Epistemology of Physics." **Berkeley Cognitive Science Report**, 48: 1-69.

Driver, R. (1983). **The Pupil as Scientist**. Milton Keynes, England and Philadelphia, PA: The Open University Press.

Driver, R., E. Guesne, and A. Tiberghien. (1985). **Children's Ideas in Science**. Philadelphia, PA: The Open University Press.

Gilbert, J., and D.M. Watts. (1983). "Concepts, Misconceptions and Alternative Conceptions: Changing Perspectives in Science Education." **Studies in Science Education**, 10: 61-98.

Goldberg, F., and L. McDermott. (February 1987). "An Investigation of Student Understanding of the Real Image Formed by a Converging Lens or Concave Mirror." **American Journal of Physics**, 55: 108-119.

Halloun, I., and D. Hestenes. (November 1985a). "The Initial Knowledge State of College Physics Students." **American Journal of Physics**, 53: 455-462.

Halloun, I., and D. Hestenes. (November 1985b). "Common Sense Concepts about Motion." **American Journal of Physics**, 53: 1056-1065.

Jung, H. (1984). "Preinstructional Conceptual Frameworks in Elementary Mechanics and Their Interaction with Instruction." Doctoral diss., University of Washington, Seattle.

Loftus, E. (1980). **Memory**. Reading, MA: Addison-Wesley.

McCloskey, M. (April 1984). "Cartoon Physics." **Psychology Today**. 52-58.

McDermott, L. (July 1984). "Research on Conceptual Understanding in Mechanics." **Physics Today**, 37: 24-32.

Minstrell, J. (1982a). "Conceptual Development Research in the Natural Setting of the Classroom." In **Education for the 80's: Science**, edited by M.B. Rowe. Washington, DC: National Education Association.

Minstrell, J. (January 1982b). "Explaining the 'At Rest' Condition of an Object." **The Physics Teacher**, 20: 10-14.

Minstrell, J. (1984). "Teaching for the Development of Understanding of Ideas: Focus on Moving Objects." In **Observing Science Classrooms: Observing Science Perspectives from Research and Practice. 1984 AETS Yearbook**. Columbus, OH: Ohio State University.

Minstrell, J., and V. Stimpson. (December 1986). "Instruction for Understanding: A Cognitive Process Framework." A Final report to NIE. NIEG83-0058.

Minstrell, J. (April 1987). "Classroom Dialogues for Promoting Physics Understanding." Paper prepared for symposium titled "The Goal of Understanding: Current Trends and Needs in Physics Education Research." AERA, Washington, DC.

Osborne, R. (November 1984). "Children's Dynamics." **The Physics Teacher**, 504-508.

Osborne, R., and P. Freyberg. (1985). **Learning in Science: The Implications of Children's Science**. Portsmouth, NH: Hinemann.

Pea, R. (1988). **Technology in Education: Looking Toward 2020**. Hillsdale, NJ: Erlbaum.

Posner, G., K. Strike, P. Hewson, and W. Gertzog. (1982). "Accommodation of a Scientific Conception: Toward a Theory of Conceptual Change." **Science Education**, 66: 211-228.

Rowe, M. (1974). "Wait-Time and Rewards as Instructional Variables: Their Influence on Language, Logic, and Fate Control." **Journal of Research in Science Teaching**, 11: 81-94.

Viennot, L. (1979). "Spontaneous Reasoning in Elementary Dynamics." **European Journal of Science Education**, 1: 205-221.

ACCOMMODATION OF A SCIENTIFIC CONCEPTION: TOWARD A THEORY OF CONCEPTUAL CHANGE

George J. Posner, Kenneth A. Strike, Peter W. Hewson, and William A. Gertzog

It has become a commonplace belief that learning is the result of the interaction between what the student is taught and his current ideas or concepts.[1] This is by no means a new view of learning. Its roots can be traced back to early Gestalt psychologists. However, Piaget's (1929, 1930) early studies of children's explanations of natural phenomena and his more recent studies of causality (Piaget, 1974) have perhaps had the greatest impact on the study of the interpretive frameworks students bring to learning situations.

This research has led to the widespread study of students' scientific misconceptions.[2] From these studies and, particularly, from recent work by researchers such as Viennot (1979) and Driver (1973), we have developed a more detailed understanding of some of these misconceptions and, more importantly, why they are so "highly robust" and typically outlive teaching which contradicts them (Viennot, 1979, p. 205).

But identifying misconceptions or, more broadly speaking, "alternative frameworks" (Driver & Easley, 1978), and understanding some reasons for their persistence, falls short of developing a reasonable view of *how* a student's current ideas interact with new, incompatible ideas. Although Piaget (1974) developed one such theory, there appears to be a need for work which focuses "more on the actual content of the pupil's ideas and less on the supposed underlying logical structures" (Driver & Easley, 1978, p. 76). Several research studies have been performed (Nussbaum, 1979; Nussbaum & Novak, 1976; Driver, 1973; Erickson, 1979) which have investigated "the *substance* of the actual beliefs and concepts held by children" (Erickson, 1979, p. 221). However, there has been no well-articulated theory explaining or describing the substantive dimensions of the process by which people's central, organizing concepts change from one set of concepts to another set, incompatible with the first. We believe that a major

---

*This article first appeared in **Science Education** (1982), Volume 66, Number 2, pp. 211-227. Copyright © 1982 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

source of hypotheses concerning this issue is contemporary philosophy of science, since a central question of recent philosophy of science is how concepts change under the impact of new ideas or new information. In this article we first sketch a general model of conceptual change which is largely derived from current philosophy of science, but which we believe can illuminate learning as well. We then illustrate some features of this model from interviews with students studying special relativity in physics. Finally, we derive some pedagogical implications.

Our central commitment in this study is that learning is a rational activity. That is, learning is fundamentally coming to comprehend and accept ideas because they are seen as intelligible and rational. Learning is thus a kind of inquiry. The student must make judgments on the basis of available evidence. It does not, of course, follow that motivational or affective variables are unimportant to the learning process. The claim that learning is a rational activity is meant to focus attention on what learning is, not what learning depends on. Learning is concerned with ideas, their structure and the evidence for them. It is not simply the acquisition of a set of correct responses, a verbal repertoire or a set of behaviors. We believe it follows that learning, like inquiry, is best viewed as a process of conceptual change. The basic question concerns how students' conceptions change under the impact of new ideas and new evidence.

The Epistemological Base

Contemporary views in philosophy of science suggest that there are two distinguishable phases of conceptual change in science. Usually scientific work is done against the background of central commitments which organize research. These central commitments define problems, indicate strategies for dealing with them, and specify criteria for what counts as solutions. Thomas Kuhn (1970) calls these central commitments[3] "paradigms," and paradigm-dominated research "normal science." Irme Lakatos (1970) labels scientists' central commitments as their "theoretical hard core" and suggests that these commitments generate "research programs" designed to apply them to and defend them from experience.

The second phase of conceptual change occurs when these central commitments require modification. Here the scientist is faced with a challenge to his basic assumptions. If inquiry is to proceed, the scientist must acquire new concepts and a new way of seeing the world. Kuhn terms this kind of conceptual change a "scientific revolution." For Lakatos it is a change of research programs.

We believe there are analogous patterns of conceptual change in learning. Sometimes students use existing concepts to deal with new phenomena. This variant of the first phase of conceptual change we call *assimilation*. Often, however, the students' current concepts are inadequate to allow him to grasp some new phenomenon successfully. Then the student must replace or reorganize his central concepts. This more radical form of conceptual change we call *accommodation*.[4]

This view of inquiry and learning involves one additional feature. We believe that inquiry and learning occur against the background of the learner's current concepts. Whenever the learner encounters a new phenomenon, he must rely on his current concepts to organize his investigation. Without such concepts it is impossible for the learner to ask a question about the phenomenon, to know what would count as an answer to the question, or to distinguish relevant from irrelevant features of the phenomenon. Without concepts the world is and remains William James' "blooming buzzing confusion." Borrowing a phrase from Stephen Toulmin (1972), we refer to those concepts which govern a conceptual change as a conceptual ecology."[5]

Our work has focused on the kinds of radical conceptual changes which we describe as accommodations. How do accommodations take place? Recent work in philosophy of science suggests that this question be divided into two. The first concerns the conditions under which an accommodation is likely to take place. When will individuals find it reasonable to undertake a major reorganization of their current concepts or to replace one set of central concepts with another? Even in a major conceptual reorganization, however, not all concepts are replaced. Individuals will retain many of their current concepts, some of which will function to guide the process of conceptual change. One can then ask what kinds of concepts tend to govern the process of accommodation. This is, in effect, to ask for the features of the conceptual ecologies which govern the process of major conceptual changes. We thus express our theory of accommodation in response to two questions:

1) Under what conditions does one central concept come to be replaced by another?

2) What are the features of a conceptual ecology which govern the selection of new concepts?

Conditions of Accommodation

The views of science on which this work is based differ from their empiricist predecessors in ways that are suggestive concerning the conditions of accommodation. Most varieties of empiricism tend to see the grounds for accepting a given scientific theory as the capacity of the theory to generate confirmed predictions. More recent views, however, suggest that an adequate view of the grounds for accepting a new theory must take into account the character of the problems generated by its predecessor and the nature of the new theory's competition.

One rather common theme in recent literature is that central concepts rarely directly entail anything about experience. Rather they suggest strategies and procedures whereby phenomena may be assimilated. Central concepts are thus not judged in terms of their immediate capacity to generate correct predictions. They are judged in terms of their resources for solving current problems. In Lakatos' terms (1970) research programs are not confirmed or refuted. Instead they are progressive or degenerative. Central concepts are likely to be rejected when they have generated a class of problems which they appear to lack the capacity to solve. A competing view will be accepted when it appears to have the potential to solve these problems and to generate a fruitful line of further research.

It is also important to note that a person's central concepts are the vehicle whereby a given range of phenomena become intelligible. Such concepts can be linked to prior experience, images, or models which make them appear intuitively obvious and which make competing concepts seem not just wrong but virtually unintelligible. Often, therefore, the first hurdle a set of central concepts must face in gaining acceptance is to appear to make sense.

These kinds of considerations suggest that there are several important conditions which must be fulfilled before an accommodation is likely to occur. The following four seem to us to express conditions which are common to most cases of accommodation.

1) *There must be dissatisfaction with existing conceptions.* Scientists and students are unlikely to make major changes in their concepts until they believe that less radical changes will not work. Thus, before an accommodation will occur, it is reasonable to suppose that an individual must have collected a store of unsolved puzzles or anomalies and lost faith in the capacity of his current concepts to solve these problems.[6]

2) *A new conception must be intelligible.* The individual must be able to grasp how experience can be structured by a new concept sufficiently to explore the possibilities inherent in it. Writers often stress the importance of analogies and metaphors in lending initial meaning and intelligibility to new concepts (Ortony, 1975; Belth, 1977; Black, 1962).

3) *A new conception must appear initially plausible.* Any new concept adopted must at least appear to have the capacity to solve the problems generated by its predecessors. Otherwise it will not appear a plausible choice. Plausibility is also a result of consistency of the concepts with other knowledge. A new idea in, say, astronomy is less likely to be accepted if it is inconsistent with current physical knowledge or if it simply has no clear physical account. Physical scientists prior to the 20th century, for example, were reluctant to accept what geologists were claiming about the age of the world since they had no theory which would allow the sum to provide energy for that period of time.

4) *A new concept should suggest the possibility of a fruitful research program.* It should have the potential to be extended, to open up new areas of inquiry.

Features of a Conceptual Ecology

An individual's current concepts, his conceptual ecology, will influence the selection of a new central concept. The literature in philosophy of science and our own work (to be discussed shortly) have suggested that the following kinds of concepts are particularly important determinants of the direction of an accommodation.

1) *Anomalies:* The character of the specific failures of a given idea are an important part of the ecology which selects its successor.

2) *Analogies and metaphors:* These can serve to suggest new ideas and to make them intelligible.

3) *Epistemological commitments:*

a) Explanatory ideals: Most fields have some subject matter-specific views concerning what counts as a successful explanation in the field.

b) General views about the character of knowledge: Some standards for successful knowledge such as elegance, economy, parsimony, and not being *ad hoc* seem subject matter neutral.

4) *Metaphysical beliefs and concepts:*

a) Metaphysical beliefs about science: Beliefs concerning the extent of orderliness, symmetry, or nonrandomness of the universe are often important in scientific work and can result in epistemological views which in turn can select or reject particular kinds of explanations. Such beliefs played a large role in Einstein's thought. Beliefs about the relations between science and commonplace experience are also important here.

b) Metaphysical concepts of science: Specific scientific concepts often have a metaphysical quality in that they are beliefs about the ultimate nature of the universe and are immune from direct empirical refutation. A belief in absolute space or time is an example.

5) *Other knowledge:*

a) Knowledge in other fields.

b) Competing concepts: One condition for the selection of a new concept is that it should appear to have more promise that its competitors.

We will see in this study how these five features of a conceptual ecology relate to the four conditions of a conceptual change in accounting for the difficulties students face in learning science. We thus turn to a study of the conceptual change required of physics students in the context of a specific topic: Einstein's special theory of relativity. This topic was chosen because it has been commonly viewed as a prototype of a scientific revolution.

The Method

In order to study students' attempts in coming to terms with the special theory, we conducted interviews in a noncalculus, self-study, self-paced introductory college physics course with students who had completed a unit on special relativity, and with several physics instructors. In the interviews, two problems were presented to the interviewees who were requested to solve the problems while thinking aloud. At each stage they were asked to give reasons for their answers, but no attempts were made to teach them in those cases where their answers were inconsistent with the special theory. The first problem considered the workings of a light clock and the implications it has for the concept of time. The second problem involved simultaneity and the synchronization of distant clocks and was followed by the presentation of written explanations from two different points of view which the interviewee was asked to read, and subsequently to repeat back, as a comprehension exercise.[7]

The Theory with Illustrations from Interviews

We now explore in greater detail the conditions governing assimilation and accommodation by relating them to the different features of a conceptual ecology listed above. The application to the special theory of relativity is illuminated by examples taken from the interviews with physics students and instructors.

Intelligibility of a New Conception

In order for a student to consider an alternative conception, he must find it intelligible. It should be clear that intelligibility is necessary for but not equivalent to or sufficient for accommodation. Intelligibility at a superficial level requires an understanding of the component terms and symbols used and the syntax of the mode of expression. For some new conceptions, this aspect of intelligibility is easily met. The special theory is one such case in which this aspect is not particularly problematic for college students with an adequate background in algebra.

However, as recent research on language comprehension demonstrates, finding discourse (or for that matter, theories) intelligible requires more than just knowing what the words and symbols mean. Intelligibility also requires constructing or identifying a coherent representation of what a passage or theory is saying (Bransford & Johnson, 1973). In fact, we would claim that no theory can function psychologically at all unless it is internally represented by the individual.

In general, representations may be in the form of propositions or images, or networks of interrelated propositions and/or images. One might, for example, represent travel distances between New York State cities as a series: Ithaca-Albany, 165 miles; Albany-Syracuse, 60 miles; Albany-New York City, 150 miles, etc. Or, the same information could be represented by a matrix, where each matrix cell contains the distance between the two cities intersecting at the cell. Or, the same information could be represented even more economically by a New York State map with lines connecting pairs of cities and distance written on each line. Similarly, a truth table and a Venn diagram can represent the same information as propositions and as images, respectively.[8]

Representations function both passively and actively. They function passively as a format into which information must be fit. In paragraph comprehension tasks, for example, anomalous sentences are confusing (i.e., unintelligible) because they cannot be fit into the representations being built and, thus, are not easily entered into the reader's memory (Bransford & Johnson, 1973). Representations also function actively as a plan for directing one's attention and conducting purposeful searches (Neisser, 1976). The inability of readers to remember an anomalous sentence in an otherwise coherent paragraph may be attributed to the readers' inattention to it.

The different functions of a representation showed up dramatically in the comparison between the responses of a student, ON, facing relativity for the first time, and an instructor, ET, who had taught physics, but not relativity, for a number of years. Both read the written explanation of the simultaneity problem, part of which follows:

> When E passed A, they both set their clocks to zero and sent me a synchronization signal. Since I was a distance d_{AB}/c from A when the signal arrived I set my clock to read $t = d_{AB}/c$. So E's clock and my clock were correctly synchronized when E was at A, but since moving clocks run slow, when E reached me, his clock had fallen behind mine.

ON's recollection of this section came without hesitation as follows:

> (B said) that when E reached point A, they synchronized the clocks to read zero, and at that point he sends a signal to B, and B synchronized his clock with *t* equal to the distance AB over the velocity and that's how B stated that he did it—because moving clocks run slow, E was behind.

There was no indication of any active functioning. She was intent on giving the written explanation without comment. She had already attempted her own solution to the problem, but there is no comparison between the written explanation and her own previous attempt. This, in contrast, is ET's recollection of precisely the same section:

> (E said) that he sent a synchronization signal to B, and that differs from approach which was to have B look at them visually. An apparently here's B considering that (E) sent out a synchronization signal and that the distance that it travelled (pause) yes, see, that's where he's getting his velocity, relative velocity.

The representation which ET builds, functions actively to direct his attention in commenting on the written explanation and to conduct a search for any information which could be used to clear up difficulties in his own solution to the problem. As he says later:

> Look, I don't remember what he said, I wasn't really trying to recall it, but to sort out my own ideas... I suppose I was selective in my reading, trying only to take from (the written explanation) what would clarify my own ideas.

How one represents knowledge and theories determines one's ability to make sense of and use the new ideas. Only if the student can psychologically construct a coherent, meaningful representation of a theory can it become an object of assessment and a tool of thought., Only an intelligible theory can be a candidate for a new conception in a conceptual change.

How difficult is this task for special relativity? Einstein (1954) describes the two basic postulates of special relativity as follows:

> ...Every universal law of nature which is valid in relation to a coordinate system C, must also be valid, as it stands, in relation to a coordinate system C', which is in uniform translatory motion relatively to C...

The second principle, on which the special theory of relativity rests, is the "principle of the constant velocity of light *in vacuo*." This principle asserts that light *in vacuo* always has a definite velocity of propagation (independent of the state of motion of the observer or of the source of the light) (pp. 224-225).

Constructing a coherent representation of the theory's two postulates individually is not particularly problematic. One can imagine a state of affairs in which each in turn is true, although the more one accepts Newtonian mechanics the harder it will be to imagine a world in which the postulate about constancy of the speed of light is true. But in balance, intelligibility of each of Einstein's two postulates is not particularly problematic.

The intelligibility of the theory as a whole, however, is a different matter. Finding it intelligible entails imagining a world in which both of Einstein's postulates are true, together with the logical implications of the postulates for notions of space and time. This task is a demanding one. To make matters even more difficult, it is possible to apply the postulates and formulas of special relativity in a superficial way without those necessary revisions in one's conceptions of space and time which are in accord with the theory; or without even having understood the full implications of its principles. Thus, both learner and instructor can mistake the intelligibility of the parts—the postulates of the special theory—for the intelligibility of the whole.

Initial Plausibility of a New Conception

One source of difficulty in learning special relativity stems from its lack of initial plausibility to physics students. Regardless of how intelligible one finds the theory, it may still appear counterintuitive. What makes a theory like special relativity counterintuitive?

Initial plausibility can be thought of as the anticipated degree of fit of a new conception into an existing conceptual ecology. There appear to be at least five ways by which a conception can become initially plausible.

1) One finds it consistent with one's current metaphysical beliefs and epistemological commitments, i.e., one's fundamental assumptions.

2) One finds the conception to be consistent with other theories or knowledge.

3) One finds the conception to be consistent with past experience.

4) One finds or can create images for the conception, which match one's sense of what the world is or could be like.

5) One finds the new conception capable of solving problems of which one is aware (i.e., resolving anomalies).[9]

Of these five factors the first appears to offer the greatest explanatory power with regard to the difficulties faced by students attempting to learn Einstein's special theory. Let us then look at fundamental assumptions as they bear on this learning task.

One set of fundamental assumptions is the individuals' epistemological commitments. Einstein (1949) was committed to two fundamental epistemological principles:

1) A theory must not contradict empirical facts; and

2) The premises of the theory must be characterized by "naturalness" or "logical simplicity," a kind of "inner perfection" of the theory. He was committed so fully to these two principles that he was able to apply them ruthlessly, even if that application meant a rejection of our common sense notions of space and time.

Needless to say, students do not always share Einstein's epistemological commitments, but their own commitments are likely to be highly significant in determining what they find initially plausible and, thus, in shaping their conceptual changes. Therefore, it is important to find out just what epistemological commitments students have, if one wants to understand what they are likely to find initially plausible or implausible and more generally, to understand their processes of conceptual change. What is their theory of theories? What is their theory of knowing? What is their view of the relation of disciplinary knowledge to everyday knowledge?

Scientific metaphysical beliefs, like epistemological commitments, are central to a conception. The central scientific metaphysical belief that contrasts special relativity with its immediate predecessors is its rejection of an absolute space and absolute time in favor of an interpretation which considers space and time relative to any given inertial system. So long as students are firmly committed to absolute space and time, they will find the special theory counterintuitive.

A good example of such a commitment is provided by student CP, who outlines her belief in absolute time explicitly and repeatedly. In response to a portion of the simultaneity problem (for which the special theory predicts that two clocks read different times) she responds:

> (CP) I mean, how could they change? Time only goes at one rate, right?

After she has read the written explanation showing the derivation of the relativistic prediction, the interviewer (I) questions her further:

> (I) And so what about this idea of absolute time?
>
> (CP) I can't say that's *not* true...

CP not only states her belief in absolute time but at a later stage shows how she defends it in a discussion of the troublesome results of the simultaneity problem which she assimilates into an existing conception:

> (I) I'm just asking you what you feel about results like that.
>
> (CP) Yeah, I mean, absolute time, it just seems to go on at a certain rate everywhere. It just seems natural that it's constant everywhere. I mean, even though you see these results.
>
> (I)...it seems these are strange results. What attitude do you take of these results?
>
> (CP) I say they don't really mean all that much; it just depends on what your frame is. It's sort of like potential energy depends on the way you define zero to be?
>
> (I) The amount of potential energy you've got?
>
> (CP) Right, all relative to what's going on.

CP's reference to potential energy is significant in pinpointing a conception which enables her to regard the values given to a variable as arbitrary, being dependent solely on the observer's point of view. She attempts to resolve some counterintuitive results of Einstein's view of time by drawing an analogy between time and potential energy. No matter that the analogy might break down with further analysis—it serves her belief in absolute time.

It is appropriate at this point to note the importance of the strength and depth of a metaphysical belief in determining whether assimilation or accommodation occurs. Because CP's commitment to absolute time is so strong, accommodation is a less attractive option than assimilation, and as a result she needs to be able to make her belief in absolute time and her understanding of special relativity consistent. She succeeds, to her satisfaction, by using the potential energy analogy.

A tutor in the course, SL, provides another example of an attempt to assimilate the findings of the special theory into an existing conception, in this case in a rather more sophisticated and detailed fashion. He shows a firm Newtonian commitment to a mechanistic view of the world which requires that objects have fixed properties such as length, mass, etc., and that explanations of phenomena should be given in terms of these objects and their interactions. In talking about the question of shrinking rods and slowing clocks, he says:

> (SL) I see them as being—as changing their length, or changing their time. But I can't talk to the person who's moving at the same velocity as the stick and the clock. He's telling me that they don't change...I feel they haven't changed, but the way I'm looking at them has changed...I guess I'm allowing for the fact that person who's seeing these things at rest, who has his clock at rest, his meter stick at rest, has (pause) a little more right to say what is really happening to the sticks.

A little later he continues:

> (SL) But I'm not at all uncomfortable with the idea of foreshortening. I do say, I do feel it is a perception. I will say it is a shortening. I know in the back of my mind that my friend who's riding along with that meter stick is telling me all the time that as far as he can tell, it's the same length and I believe what he's saying, which is o.k.
>
> (I) It's not a conflict?

> (SL) No, because the fact that it's moving makes it appear to me as if it were foreshortened.

Here SL insists on treating length as a constant, independent of frames of reference. He is, thus, led to treat the special theory's claims concerning the relativity of length as simply a distortion of perception.

What is of interest to us at this point is that SL reveals this commitment by using it as the conception to which he assimilates the findings of special relativity. In order to do this he has to make two auxiliary assumptions: that a shrinking rod constitutes a perceptual problem, and doesn't actually shrink ("I feel they [rods and clocks] haven't changed, but the way I'm looking at them has changed"), and that in principle a mechanistic interpretation in terms of objects and their motion is needed in order to explain why clocks run slower ("I don't see how in depth...but I believe it can be done"). Neither of these assumptions is necessary or even consistent with an Einsteinian perspective based on a reanalysis of space and time. They do, however, play an integral part in protecting SL's metaphysical commitments.

Dissatisfaction with Existing Conceptions[10]

Generally, a new conception is unlikely to displace an old one, unless the old one encounters difficulties,[11] and a new intelligible and initially plausible conception is available that resolves these difficulties. That is, the individual must first view an existing conception with some dissatisfaction before he will seriously consider a new one.

One major source of dissatisfaction is the anomaly. Each time a person unsuccessfully attempts to assimilate an experience or a new conception into his existing network of conceptions, that person experiences an anomaly. An anomaly exists when one is unable to assimilate something that is presumed assimilable—or (in other words), one simply cannot make sense of something.

When faced with an anomaly, the individual (scientist or student) has several alternatives. One may come to the conclusion that one's existing conceptions require some fundamental revisions (i.e., an accommodation) in order to eliminate the conflict. But this is the most difficult and, therefore, the most unlikely approach, especially when there are other possibilities:

1) rejection of the observational theory[12];

2) a lack of concern with experimental findings on the grounds that they are irrelevant to one's current conception;

3) a compartmentalization of knowledge to prevent the new information from conflicting with existing belief ("Science doesn't have anything to do with the 'real' world"); and

4) an attempt to assimilate the new information into existing conceptions (e.g. "Newtonizing" relativistic phenomena).

This analysis suggests that the presentation of anomalies will produce dissatisfaction with an existing conception only if:

1) Students understand why the experimental finding represents an anomaly;

2) Students believe that it is necessary to reconcile the findings with their existing conceptions;

> 3) Students are committed to the reduction of inconsistencies among the beliefs they hold; and
>
> 4) Attempts to assimilate the findings into the students' existing conceptions are seen not to work.

Given the improbability that all these conditions will be met, it is no wonder that few students find their current conceptions weakened by anomalies. Why consider alternatives to a Newtonian view (or whatever view they hold) when they are unconvinced of the inadequacy of their conceptions? The search for instructionally viable and effective anomalies is of primary importance if accommodation is to be taken seriously as a goal. Recounting historical anomalies (such as the Michelson-Morley experiment) won't always do!

Assuming this formidable instructional problem is solved, the process of accommodation can proceed. If the dissatisfaction with the existing conception created by its inability to make sense of experience is followed by learning of an intelligible alternative which resolves or promises to resolve some of the anomalies of its predecessor, then the new conception may be plausible.

There is little evidence in the interviews that students were aware of anomalies, even though videotapes of two experiments formed part of the study material in the course. It is significant that the clearest example of a student's awareness of anomalous behavior occurs along with a statement of his epistemological commitment.

HU has derived a mistaken result from his view of the relativity principle which implies that pictures taken by two cameras moving past one another at the same instant and of the same two clocks will show different things. The interviewer looks for confirmation of this view:

> (I) So what you're saying is that they wouldn't agree, they couldn't agree, that they'd actually see different things.
>
> (HU) Right.
>
> (I) That doesn't bother you?
>
> (HU) It did at first, but when you think about it and hash it out, there really is no reason why we should limit ourselves to one frame of mind. I like to think abstractly and I can see that. I had trouble realizing that lengths would change, too, but you know, I'm game! No, it doesn't bother me. It's just that we don't realize it due to our slow speeds. I tend to agree with scientific data that's brought up and when they say that an electron—what was that—a meson, actually goes with the predictions, what can you do? And once you see the facts, you can stretch your imagination.

HU's stated epistemology is simple and empiricist: Theories are derived from experimental evidence. It also appears to exhibit a degree of tolerance for theoretical inconsistencies, which tolerance precludes him from seeing that he has made a mistake. However, HU has pinpointed one anomaly: Newtonian mechanics calculates the lifetime of a meson to be much shorter than that which is observed experimentally. He has, however, seen that special relativity's prediction agrees with experiment. Thus, HU sees an anomaly, he sees the alternative conception, and his epistemological commitment allows for its plausibility.

Fruitfulness of a New Conception

Once aware of an intelligible, plausible alternative to an existing conception that resolves apparent anomalies, students may actively attempt to map their new conceptions onto the world; that is, they may attempt to interpret experience with it. If the new conception not only resolves its predecessor's anomalies but also leads to new insights and discoveries, then the new conception will appear fruitful and the accommodation of it will seem persuasive.

A brief examination of the fruitfulness of the special theory for professional scientists may suggest some of the theory's potential. It is this potential which students should be made aware, if they are to share in the view that the theory is indeed fruitful, and, thus, worth accommodating:

1) As an engineering tool in the design of accelerators (relativistic mechanics);

2) As a technological tool in the development of nuclear weapons and nuclear reactors;

3) As a theoretical and technological tool in nuclear chemistry for predicting the products of nuclear reactions;

4) As a theoretical and mathematical tool in astronomy for calculating life expectancies of stars, for explaining astronomical phenomena (e.g. supernovae), and for making calculations of astronomical distances;

5) As a theoretical basis (along with quantum physics) for the development of modern physics.

To the extent that the student can understand these contributions and trace them back to the basic postulates of the special theory, they may begin to appreciate the fruitfulness of the theory.

The Character of Accommodation

Our description of the four conditions of a successful accommodation may have suggested a fairly straightforward linear process: students' dissatisfaction with Newtonian physics; followed by the students' finding special relativity intelligible; leading to an initial belief in its plausibility; and concluding with the belief that the theory is ultimately fruitful.

However, it should be clear that this account is oversimplified, since many basic conceptions, including relativity, are so complex that at a particular time one is likely to accommodate certain aspects but not others. We have, of course, described accommodation as a radical change in a person's conceptual system. That an accommodation is a radical change does not, however, entail that it is abrupt. Indeed, there are good reasons to suppose that for students accommodation will be a gradual and piecemeal affair. Students are unlikely to have at the outset a clear or well-developed grasp of any given theory and what it entails about the world. For them, accommodation may be a process of taking an initial step toward a new conception by accepting some of its claims and then gradually modifying other ideas, as they more fully realize the meaning and implication of these new commitments. Accommodation, particularly for the novice, is best thought of as a gradual adjustment in one's conception, each new adjustment laying the groundwork for further adjustments but where the end result is a substantial reorganization or change in one's central concepts.

Our interviews also indicate that what may initially appear as an accommodation may turn out to be something less than that. As the interview with SL indicates, people who accept Einstein's two postulates may understand them in a rather non-Einsteinian fashion. Often it appears that as students, who have accepted the two postulates, begin to realize their counterintuitive implications or their conflicts with Newtonian notions of space and time, the commitment to the two postulates weakens. Typically, students will attempt various strategies to escape the full implication of the two postulates or to reconcile them with Newtonian assumptions. Accommodation may, thus, have to wait until some unfruitful attempts at assimilation are worked through. It rarely seems characterized by either a flash of insight, in which old ideas fall away to be replaced by new visions, or as a steady logical progression from one commitment to another. Rather, it involves much fumbling about, many false starts and mistakes, and frequent reversals of direction.

Educational Implications

Teaching science involves providing a rational basis for a conceptual change. We have also seen that fundamental conceptual changes, termed accommodations, may involve changes in one's fundamental assumptions about the world, about knowledge, and about knowing and that such changes can be strenuous and potentially threatening, particularly when the individual is firmly committed to prior assumptions. We have seen that people resist making such changes, unless they are dissatisfied with their current concepts and find an intelligible and plausible alternative that appears fruitful for further inquiry.

Two features of a conceptual ecology, in particular, were shown to guide the change process from one conception to another: 1) anomalies, and 2) fundamental assumptions about science and about knowledge.

If taken seriously by students, anomalies provide the sort of cognitive conflict (like a Kuhnian state of "crisis") that prepares the student's conceptual ecology for an accommodation. The more students consider the anomaly to be serious, the more dissatisfied they will be with current concepts, and the more likely they may be ready ultimately to accommodate new ones.

Metaphysical beliefs and epistemological commitments form the basis on which judgments are made about new knowledge. Thus, a conceptual change will be rational to the extent that students have at their disposal the requisite standards of judgment necessary for the change. If a change to special relativity requires a commitment to the parsimony and symmetry of physical theories (as it did for Einstein), then students without these commitments will have no rational basis for such a change. Faced with such a situation students, if they are to accept the theory, will be forced to do so on non-rational bases, for example, because the book or the instructors says it is "true."

Our study of the history of science reveals that many conceptual changes in science have been driven by the scientists' fundamental assumptions rather than by the awareness of empirical anomalies.[13] Einstein's special relativity can be seen as such a case.[14] However, since it is unlikely that students in an introductory physics course can be successfully taught the requisite standards of judgment for an accommodation of special relativity, physics teachers must rely on "anomalies" to prepare the student for the accommodation.

Our problem remains unsolved, however. Most of the anomalies will not be readily seen as anomalies by students without a thorough understanding of the observational theory in which the experiment was embedded. That is, most of the experiments are far from being "transparent." Does this problem mean that the special theory can realistically be made at best only intelligible and partially plausible, but never fully persuasive to students who are firmly committed to a set of conflicting metaphysical beliefs and epistemological commitments? It is one thing to educate physicists over a course of four to six years into a given set of standards of judgment. It is quite another thing to accomplish this goal in an introductory physics course, along with covering a great deal of content. And further, how is one to demonstrate the theory's fruitfulness in the limited time available?

Accepting, then, that accommodation of the special theory runs the risk of being difficult, if not impossible, is there anything we as educators can do to enable physics students to accommodate new conceptions on a rational basis? Let us examine the implications of our research for science education. We shall frame these implications in the form of questions and suggestions raised by our research thus far.

Curricular Objectives

Our discussion of the critical role played by the student's fundamental assumptions about the world and about their knowledge of the world raises serious questions about the objectives of science courses. If the conceptual change process is to be rationally based, then students will need to be immunized against the kind of inevitable indoctrination that occurs when neither the teacher nor the student is aware of his own fundamental assumptions, much less those implied by the science he is teaching and learning.

The primary question which must be raised about curriculum objectives as a result of our discussion in this article is the following: Is it realistic to expect science instruction to produce accommodation in students, rather than merely to help students make sense of new theories? And secondarily, should this be an expectation for all students, or only for certain groups, such as science majors?

In the event that an affirmative answer is given to the primary question, the contents of the previous sections suggest we aim at developing in students:

1) An awareness of their fundamental assumptions and of those implicit in scientific theory;

2) A demand for consistency among their beliefs about the world;

3) An awareness of the epistemological and historical foundations of modern science;

4) Some sense of the fruitfulness of new conceptions.

The extent to which any of the above should be considered is a matter for future investigation.

Content

If we aim to produce rationally based conceptual change in students, then according to what we have said thus far, the content of science courses should be such that it renders scientific theory intelligible, plausible, and fruitful. In order to give expression to this general requirement, the following conditions appear to be necessary:

1) More emphasis should be given to assimilation and accommodation by students of that content than to content "coverage."

2) "Retrospective anomalies" should be included, particularly if historically valid anomalies are difficult to comprehend, or, as with the special theory, were not responsible for driving the conceptual change in the first place.[15]

3) Sufficient observational theory should be taught for students to understand the anomalies employed.

4) Any available metaphors, models, and analogies should be used to make a new conception more intelligible and plausible.

Teaching Strategies

Teaching is typically thought of as clarifying content presented in texts, explaining solutions to problems, demonstrating principles, providing laboratory exercises, and testing for recall of facts and ability to apply knowledge to problems. That is, teaching is for recall and assimilation. For teaching aimed at accommodation the following possible changes in this approach are implied by our research:

1) Develop lectures, demonstrations, problems, and labs which can be used to create cognitive conflicts in students. Among other things, one might consider what types of homework problems would create the kind of cognitive conflict necessary as preparation for an accommodation, and whether labs could be used to help students experience anomalies (Stavy & Berkowitz, 1980).

2) Organize instruction so that teachers can spend a substantial portion of their time in diagnosing errors in student thinking and identifying defensive moves used by students to resist accommodation.

3) Develop the kinds of strategies which teachers could include in their repertoire to deal with student errors and moves that interfere with accommodation.

4) Help students make sense of science content by representing content in multiple modes (e.g., verbal, mathematical, concrete-practical, pictorial), and by helping students translate from one mode of representation to another (Clement, 1977).

5) Develop evaluation techniques to help the teacher track the process of conceptual change in students (e.g., the Piagetian clinical interview) (Posner & Gertzog, 1982).

Teacher Role

The teacher as clarifier of ideas and presenter of information is clearly not adequate for helping students accommodate new conceptions. Our research suggests that the teacher might have to assume two further roles in order to facilitate student accommodation. In these roles the teacher would become:

1) An adversary in the sense of a Socratic tutor. In this role, the teacher confronts the students with the problem arising from their attempts to assimilate new conceptions. (A point of concern is the need to avoid establishing an adversarial role with regard to students as persons while developing and maintaining it with regard to conceptions.)

2) A model of scientific thinking. Aspects of such a model might include a ruthless demand for consistency among beliefs and between theory and empirical evidence, a

pursuit of parsimony among beliefs, a skepticism for excessive "*ad hoc*-ness" in theories and a critical appreciation of whether discrepancies between results may be in "reasonable agreement" with theory.[16]

Whether any of the above changes could be implemented, and the extent to which they would prove effective in facilitating accommodation in students, are questions which we have not as yet addressed.

Authors' Note: This article is partly based on a paper entitled "Learning Special Relativity: A Study of Intellectual Problems Faced by College Students," presented at the International Conference Celebrating the 100th Anniversary of Albert Einstein, November 8-10, 1979 at Hofstra University.

Notes

1. See, for example, Ausubel (1968).

2. See Driver and Easley (1978) for an excellent review of research in this area.

3. In this article the terms "commitments," "concepts," and "conceptions" are equivalent. Although these terms (particularly the latter two) refer to differing levels of conceptualization, we do not distinguish among the levels in this article.

4. These are Piaget's words, but in using them we do not intend any commitment to his theories.

5. We understand this view in direct opposition to traditional empiricism. Empiricism's central commitment, that there is nothing in the mind not first in the senses, requires people to be able to learn something in the total absence of any prior concepts. We believe this to be impossible. A mind which began as a blank tablet would remain so, for it would lack the means to investigate experience.

6. There is, of course, a sizeable body of literature in both psychology (Smedslund, 1961; Kuhn, 1972; Berlyne, 1965) and science education (Driver, 1973; Stavy & Berkowitz, 1980) on the use of conceptual or cognitive conflict for the development of thought and conceptual change. But most of this literature has been conducted within a strictly Piagetian framework [though Berlyne (1965) reviews the broader use of conflict situations in education]. However, none of this work appears to be grounded in a theory of conceptual change of the sort discussed in this article. That is, none is focused on fundamental changes in a person's central, organizing concepts from one set of concepts to another set incompatible with the first.

7. See Posner et al. (1979) for a full description of the problems.

8. These examples were all adapted from Jerome Bruner (1966).

9. The present discussion will focus only on the first of these five ways, that is, on the individual's fundamental assumptions. In the next section we elaborate on the last of these ways. There we discuss the dual function of anomalies.

10. The reader should note that this section builds on some of the points raised in the two previous sections. Therefore, for purposes of clarity and succinctness, this section is placed out of order in relation to the list of conditions presented earlier.

11. Lakatos (1970) terms these difficulties "recalcitrant data."

12. Reports of observation are not theory neutral. Rather, observations are described and interpreted by means of concepts taken from some theory, or some theory is assumed in treating the observations as data. Treating observations of the red shift as a measure of distance assumes a wave theory of light and, in some cases, relativity. Even the use of a telescope assumes a theory of optics. Theories which function to describe or interpret data we refer to as observation theories.

13. See Burtt's (1962, pp. 36-62) account of Copernicus, whose theory was not a response to anomalies, but was only presented as a simpler and more harmonious interpretation.

14. Empirical findings anomalous with respect to Newtonian physics but consistent with Einsteinian theory developed many years after the special theory of relativity was proposed.

15. See Anthony P. French (1968, pp. 6-29), for an example of the use of retrospective anomalies in teaching special relativity.

16. See Thomas S. Kuhn (1977, pp. 178-224), for an insightful discussion of this point. For a useful discussion of the teacher as a model of scientific thought, see Gene D'Amour (1977, pp. 183-190).

References

Ausubel, D.P. *Educational psychology: A cognitive viewpoint*. New York: Holt, Rinehart & Winston, 1968.

Belth, M. *The process of thinking*. New York: David McKay, 1977.

Berlyne, D.E. Curiosity and education. In *Learning and the educational process*, J.D. Krumboltz (Ed.). Chicago: Rand McNally, 1965.

Black, M. *Models and metaphors*. Ithaca, NY: Cornell University Press, 1962.

Bransford, J.D., & Johnson, M.K. Considerations of some problems of comprehension. In *Visual information processing*, W. G. Chase (Ed.). New York: Academic Press, 1973.

Bruner, J.S. *Toward a theory of instruction*. Cambridge: Harvard University Press, 1966.

Burtt, E.A. *The metaphysical foundations of modern science* (rev. ed.). Garden City, NY: Doubleday Anchor Books, 1932.

Clement, J.J. *Some types of knowledge used in understanding physics*. University of Massachusetts, Department of Physics and Astronomy, 1977 (mimeographed).

D'Amour, G. Problem-solving strategies and the epistemology of science. In *Cognitive process instruction*, J. Lochhead & J.J. Clement (Eds.). Philadelphia: The Franklin Press, 1979.

Driver, R. The representation of conceptual frameworks in young adolescent science students. Unpublished Ph.D. dissertation, University of Illinois, Urbana, Illinois, 1973.

Driver, R., & Easley, J. Pupils and paradigms: A review of literature related to concept development in adolescent science students. *Stud. Sci. Ed.*, 1978, *5*, 61-84.

Einstein, A. Autobiographical notes. In *Albert Einstein: Philosopher-scientist*, P.A. Schilpp (Ed.). Evanston, IL: Library of Living Philosophers, 1949.

Erickson, G.L. Children's conceptions of heat and temperature. *Sci. Educ.*, 1979, *63*, 221-230.

French, A.P. *Special relativity*. New York: W.W. Norton, 1968.

Kuhn, D. Mechanisms of change in the development of cognitive structure. *Child Dev.*, 1972, *43*, 823.

Kuhn, T.S. *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press, 1970.

Kuhn, T.S. The function of measurement in modern physical science. In *The essential tension: Selected studies in scientific tradition and change*. Chicago: University of Chicago Press, 1977.

Lakatos, I. Falsification and the methodology of scientific research programmes. In *Criticism and the growth of knowledge*, I. Lakatos & A. Musgrave (Eds.). Cambridge: Cambridge University Press, 1970.

Neisser, U. *Cognition and reality*. San Francisco: W.H. Freeman, 1976.

Nussbaum, J. Children's conception of the Earth as a cosmic body: A cross age study. *Sci. Educ.*, 1979, *63*, 83-93.

Nussbaum, J., & Novak, J.D. An assessment of children's concepts of the Earth utilizing structured interviews. *Sci. Educ.*, 1976, *60*, 535-550.

Ortony, A. Why metaphors are necessary and not just nice. *Educ. Theory*, 1975, *25*, 45-53.

Piaget, J. *Understanding causality.* New York: W.W. Norton, 1974.

Piaget, J. *The child's conception of the world.* New York: Harcourt Brace, 1929.

Piaget, J. *The child's conception of physical causality.* London: Kegan Paul, 1930.

Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. Learning special relativity: A study of intellectual problems faced by college students. Paper presented at the 100th Anniversary of the Birth of Albert Einstein, Hofstra University, 1979.

Posner, G.J., & Gertzog, W.A. The clinical interview and the measurement of conceptual change. *Sci. Ed.*, 1982, *66*, 195-209.

Smedslund, J. The acquisition of conservation of substance and weight in children: V. Practice in conflict situations without external reinforcement. *Scand. J. Psychol.*, 1961, *2*, 156-160.

Stavy, R., & Berkowitz, B. Cognitive conflict as a basis for teaching quantitative aspects of the concept of temperature. *Sci. Educ.*, 1980, *64*, 679-692.

Toulmin, S. *Human understanding.* Princeton: Princeton University Press, 1972.

Viennot, L. Spontaneous reasoning in elementary dynamics. *Eur. J. Sci. Educ.*, 1979, *1*, 205-221.

INDEX

A

ability grouping 43–68
Aldridge, Bill G. 1, 9
Anderson, Charles W. 158, 221
assessment 142
  and constructivism 181
  performance-based 16
  standardized tests 60–61, 132

C

community
  as it relates to opportunity 79
conceptual change 158, 225, 253–269, 255
  conditions of accommodation 255
  features of a conceptual ecology 256
  preparation for 230–231
constructivism 169–182, 183–194
  and problem solving 179
  and science instruction 177
  as a form of empiricism 187
  radical 170, 186
  social 170
  trivial 170
cooperative learning 61–62, 69
  and academic achievement 71
  and constructivism 176
  and effects on multiethnic relations 73
  and self-esteem 71
  and social skills 71
coordination 1
course-taking patterns 3–7, 58

D

Dempster, Frank N. 25

G

gatekeeping courses 99
Gertzog, William A. 253
girls
  and achievement 111–124
  self-confidence 116

H

hands-on science 153, 224
Hewson, Peter W. 253
higher order skills 129–144
  defined 130

I

inquiry perspective 153, 224

K

Kahle, Jean Butler 111
knowledge integration 197–218
  integrating concepts of force 243
  defined 199

L

layer cake 9
Linn, Marcia C. 111, 116, 197
Lucking, Robert 69

M

Manning, M. Lee 69
Matthews, Michael R. 183
Minstrell, James A. 200, 230, 237
misconceptions 226, 238–242, 253–254
modeling 137, 267
motivation 2, 15, 138

O

Oakes, Jeannie 79

P

Posner, George J. 158, 244, 253
problem solving 134–135. *See also*
 higher order skills
 and constructivism 179

R

race
 and access to mathematics and
  science courses 94–96, 98
 effects on opportunity 79–88
Resnick, Lauren B. 129–144
Roth, Kathleen J. 151, 221, 227

S

science process skills. *See* inquiry
 perspective
Science-Technology-Society 15, 155
scope 1
sequence 1
Slavin, Robert E. 43, 69
social class
 and access to mathematics and
  science courses 94, 98
 as it relates to opportunity 79–88
Songer, Nancy Butler 197
spacing effect 10, 25–38
streaming. *See* ability grouping
Strike, Kenneth A. 253

T

terms 11
 and constructivism 180
tracking 82. *See also* ability grouping.

V

von Glasersfeld,
 Ernst 169, 184, 186, 194

W

wait time 241, 249

SS&C Advisory Board

LaMoine L. Motz, CHAIR, Waterford, MI
Hans O. Andersen, Bloomington, IN
Jean Beard, Sunnyvale, CA
Napolean A. Bryant, Cincinnati, OH
Michael H. Farmer, Greenville, SC
Lynn W. Glass, Ames, IA
Joan Hall, Cincinnati, OH
David R. Stronck, Hayward, CA
Michael Burton, Fargo, ND
Nancy Ridenour, Ithaca, NY
Thomas Rossing, DeKalb, IL
F. James Rutherford, Washington, DC
Kendell Starkweather, Reston, VA
Robert Christman, Bellingham, WA
Sylvia Ware, Washington, DC

NSTA/SS&C Coordinating Center Staff

Arnold Strassenburg
Project Director
Russell Aiuto
Director of Research and Development
Erma Anderson
National Project Manager

Stephanie M. Jackson
Administrative Assistant
Jannine Jordan
Program Assistant

VOLUME II
RELEVANT RESEARCH

Bill Aldridge
Charles W. Anderson
Frank N. Dempster
William A. Gertzog
Peter W. Hewson
Jean Butler Kahle
Marcia C. Linn
Robert Lucking
M. Lee Manning
Michael Matthews
James A. Minstrell
Jeannie Oakes
George J. Posner
Lauren B. Resnick
Kathleen J. Roth
Nancy Butler Songer
Robert E. Slavin
Kenneth A. Strike
Ernst von Glasersfeld

VOLUME II: RELEVANT RESEARCH makes available to the science educator a select group of significant research papers and philosophical statements on how secondary school students learn science best. Its papers address current learning research issues such as spacing content (Dempster), cooperative learning (Slavin, Manning and Lucking), equitable classrooms (Oakes, Kahle) cognitive skill development (Resnick, Roth), constructivism (von Glasersfeld, Matthews, Songer and Linn) and conceptual change (Minstrell, Anderson, Posner et al.). Chapter introductions indicate connections between the research evidence and the tenets of SS&C. Taken together, these papers illuminate the multi-disciplinary research base of the Project on Scope, Sequence, and Coordination of Secondary School Science.

With its inclusion of complete bibliographies and an index of key terms in science education, this volume serves as an excellent resource for science educators. Science teachers and supervisors can use this research evidence to make decisions about the design of science curricula or courses. Additionally, it is a valuable text for teacher inservice or preservice programs.

The Project on Scope, Sequence, and Coordination of Secondary School Science (SS&C) is a major national project designed to reform science education. Based on research on learning science, the project recommends that all students study science every year for six years and advocates carefully sequenced, well-coordinated science instruction. The SS&C Project is funded by the National Science Foundation and the U.S. Department of Education.

NSTA
Washington, DC 20009

ISBN: 0-87355-111-7

BEST COPY AVAILABLE